NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

EVALUATION OF DATA PROCESSING TECHNIQUES FOR UNOBTRUSIVE GAIT AUTHENTICATION

by

William A. Parker

March 2014

Thesis Advisor:	Craig Martell
Co-Advisor:	Mark Gondree

THIS PAGE INTENTIONALLY LEFT BLANK

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE March 2014	3. REPORT TYPE AND DATES COVERED Master's Thesis
4. TITLE AND SUBTITLE EVALUATION OF DATA PROCESSING TECHNIQUES FOR UNOBTRUSIVE GAIT AUTHENTICATION		**5. FUNDING NUMBERS**
6. AUTHOR(S) William A. Parker		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943–5000		**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A		**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number ____NPS.2013.0058-IR-EP7-A____.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited	**12b. DISTRIBUTION CODE**

13. ABSTRACT (maximum 200 words)

The growth in smartphone usage has led to increased storage of sensitive data on these easily lost or stolen devices. In order to mitigate the effects of users who ignore, disable, or circumvent authentication measures like passwords, we evaluate a method employing gait as a source of identifying information.

This research is based on previously reported methods with a goal of evaluating gait signal processing and classification techniques. This thesis evaluates the performance of four signal normalization techniques (raw signal, zero-scaled, gravity-rotated, and gravity rotated with zero-scaling). Additionally, we evaluate the effect of carrying position on classification. Data was captured from 23 subjects carrying the device in the front pocket, back pocket, and on the hip. Unlike previous research, we analyzed classifier performance on data collected from multiple positions and tested on each individual location, which would be necessary in a robust, deployable system.

Our results indicate that restricting device position can achieve the best overall performance using zero-scaling with 6.13% total error rate (TER) on the XY-axis but with a high variance across different axes. Using data from all positions with gravity rotation can achieve 12.6% TER with a low statistical variance.

14. SUBJECT TERMS Smartphone authentication, accelerometer gait recognition, signal processing, quaternion rotation, gravity calibration, machine learning, support vector machine, k-nearest neighbors	**15. NUMBER OF PAGES** 87
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

THIS PAGE INTENTIONALLY LEFT BLANK

EVALUATION OF DATA PROCESSING TECHNIQUES FOR UNOBTRUSIVE GAIT AUTHENTICATION

William A. Parker
Lieutenant, United States Navy
B.S., University of North Texas, 2007

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN COMPUTER SCIENCE

from the

NAVAL POSTGRADUATE SCHOOL
March 2014

Author: William A. Parker

Approved by: Craig Martell
                    Thesis Advisor

                    Mark Gondree
                    Co-Advisor

                    Peter Denning
                    Chair, Department of Computer Science

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The growth in smartphone usage has led to increased storage of sensitive data on these easily lost or stolen devices. In order to mitigate the effects of users who ignore, disable, or circumvent authentication measures like passwords, we evaluate a method employing gait as a source of identifying information.

This research is based on previously reported methods with a goal of evaluating gait signal processing and classification techniques. This thesis evaluates the performance of four signal normalization techniques (raw signal, zero-scaled, gravity-rotated, and gravity rotated with zero-scaling). Additionally, we evaluate the effect of carrying position on classification. Data was captured from 23 subjects carrying the device in the front pocket, back pocket, and on the hip. Unlike previous research, we analyzed classifier performance on data collected from multiple positions and tested on each individual location, which would be necessary in a robust, deployable system.

Our results indicate that restricting device position can achieve the best overall performance using zero-scaling with 6.13% total error rate (TER) on the XY-axis but with a high variance across different axes. Using data from all positions with gravity rotation can achieve 12.6% TER with a low statistical variance.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

API	application programming interface
BFCC	bark-frequency cepstral coefficient
DCT	discrete cosine transform
DFT	discrete Fourier transform
DTW	dynamic time warping
EER	equal error rate
FAR	false accept rate
FFT	fast Fourier transform
FMR	false match rate
FNMR	false non-match rate
FRR	false reject rate
FS	floor sensors
FTA	failure to acquire
HTER	half total error rate
kNN	k-nearest neighbor
MFCC	mel-frequency cepstral coefficient
mTER	mean total error rate
MV	machine vision
RBF	radial basis function
SDK	software developers' kit
SVM	support vector machine
TER	total error rate
WS	wearable sensors

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

Foremost, I want to thank my mom, who always supported my academic pursuits with assurance that I was capable of anything I put my mind to. I also owe this to my brother, Jesse, who has always been a tough sibling with whom to have a rivalry, due to his constant achievements challenging me to do better.

Without my friends, here in Monterey and elsewhere, I would have overstressed myself in this endeavor. Thank you, Christine, for almost always being available to chat and distract me from otherwise over-thinking problems. I am grateful for my dog, Ozzy, who put up with late night walks while I worked on typing this thing up.

Finally, to the U.S. Navy and the American taxpayers, who put their faith in my ability to carry out my duties while funding this research and paying my bills, I owe a debt that I may never be able to repay fully.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

According to a 2011 survey conducted by Ponemon [1], companies reported approximately four percent of employee-issued smartphones were lost or stolen. While it is assessed that 60% of these lost phones contain sensitive and confidential information, 57% of them were reported to not employ data protection mechanisms. The issue of mobile phone theft has become significant enough that, in 2012, the FCC, along with leaders of major metropolitan cities, announced new initiatives to reduce theft and encourage users to better protect their data [2]. In 2012, it was reported that nearly 85% of smartphone users perform both work and personal tasks on their mobile devices [3]. Of those users who employ passcodes, two-thirds report writing their passwords down on a piece of paper, against best security practices. As such, the study and evaluation of potential methods for authorizing, and denying, data access is crucial.

The use of non-intrusive, passive authentication techniques is of particular interest. Since 2005, nearly all smartphones have contained built in tri-axial accelerometers, which can detect the rate of change in the speed of movement in lateral, longitudinal, and vertical directions. Arghire [4] predicted that one out of three mobile devices would ship with accelerometers in 2010 with an increasing trend. Additionally, the implementation of the accelerometer listener in the Android SDK does not require direct user permission or known involvement to collect and analyze data, thus creating the potential for an unobtrusive system. The goal of this research is to examine the effects of on-body placement and data normalization, while determining the most discriminatory set of variables for passive gait authentication.

A. MOTIVATION

With the increased worldwide usage of smartphones, sensitive data now travels more frequently, both electronically and physically, from work to home and places in between. This increases the chances of loss or theft of devices storing sensitive information. Ponemon [3] observed that of 116 companies surveyed, 62% of devices lost or stolen contained some sensitive data. Government agencies, including the Department

of Defense, are particularly sensitive to data leakage. The Attorney General of New York even asked Google, Apple, and other device manufacturers to take action to stop smartphone theft and ebb the increase in the black market smartphone trade [5]. In order to protect the data from potential compromise, new data access control mechanisms have been proposed.

Smartphones typically ship with built-in screen locking and PIN authentication functions, but Ponemon [3] found that 60% of the companies surveyed reported that their employees would ignore or disable security mechanisms such as passwords and keylocks. Therefore, it is critical to develop methods of authentication that do not require user attention to invoke, such as passive authentication.

Previous research by Gufarov *et al.* [6] and Nickel [7] have shown that the rhythm of an individual's walk, henceforth referred to as *gait,* can be detected by smartphone accelerometers. Further, the accelerometer signal for each individual is sufficiently characteristic that it has an acceptable recognition rate for authentication. However, these previous studies all required the device to be located at and attached to the hip, thus fixing the orientation and removing variations in signal caused by different carry positions.

In this study, we will continue to advance toward a deployable gait recognition system that can operate in real-world situations. That is, we will evaluate the effects of placing the device in different body locations, using different normalization techniques, to simulate the differences in user carrying preference.

B.   RESEARCH QUESTIONS

This thesis addresses the question of whether gait authentication methods could be improved through more discriminatory selection of data and alternate signal processing techniques. In order to address this, the following sub-questions will be evaluated.

- Would current well-performing gait authentication methods benefit from the use of an alternative classifier type?

- Can current methods be improved through the selection of alternate primary axes?

- Can the classification performance of current techniques be improved with alternative data normalization methods?

- Do current well-performing methods show similar performance when data is captured from different on-body carrying positions?

- Can an authentication system be developed that will show similar classification performance regardless of where the data capture device is carried on the body?

In order to answer these questions, we collected gait data from multiple subjects and implemented an authentication system using previously reported settings but modified to perform multiple normalization techniques on multiple classifiers. We ran a range of experiments on the system in order to determine what methods performed the best on our data set. Finally, we evaluate the performance of data from individual locations and implement a classifier that evaluates performance on data from all carrying positions.

C.   SIGNIFICANT FINDINGS

After running experiments with two classification techniques, on five axis feature sets, using four normalization techniques, from three on-body carrying positions, we report several important findings.

- Of the 60 experiments run, in 51 cases kNN classifiers showed a lower error rate when compared to SVM.

- When experimenting on individual carrying positions, we observed in the hip carrying position, the zero-scaled, combined XY-axis showed the best overall TER of 6.13%. This result was similar to the best performing features, though an improvement over the 10.7% mTER from Brandt [8]. This suggests the zero-scaling technique may be optimal when the device is located on the hip in a stable holder, which agrees with Vildjiounaite [9].

- When performing position-independent analysis, we observed that regardless of the device's carrying position, the individual Y-axis when rotated due to gravity achieved the best mTER of 12.6%. Additionally, the combined XY and XYZ axes, when rotated, achieved satisfactory mTERs of 15.4% and 18.0%, respectively, indicating the significance of the effect of gravity on the Y, or vertical axis.

- A carrying-position-independent gait authentication system can produce consistent performance results when using a gravity rotated XY-axis, and the signal processing techniques of Nickel [7] and Brandt [8].

- Normalization technique is important when training a classifier on from all known carrying positions and tested on samples from any one of the positions. The zero-scaling technique yielded better than a 30% TER in all cases, while the gravity rotation technique performed better than 20% TER with a reduced statistical variance across experiments.

- When training and testing on data from the same carrying position, the classifier performances of individual experiments are slightly worse using the gravity rotation technique than zero-scaling; however, the performance when combining all positions and testing on any one is equivalent, or better, than training and testing on the back pocket position alone. This implies that a system allowing the user to carry the device in multiple positions may perform slightly worse than position restrictive techniques, but can achieve consistent and acceptable performance when training on all positions.

D.   ORGANIZATION OF THIS THESIS

This thesis is organized as follows:

- Chapter I describes the justification for studying gait authentication techniques

- Chapter II discusses previous research in the fields of gait biometrics and smartphone accelerometer recognition techniques

- Chapter III explains of the experimental methodology and reasoning for system design decisions

- Chapter IV describes the results of the experiments, including an analysis of the findings

- Chapter V details the limitations of this work that explain the findings and offers recommendations for future research.

II. BACKGROUND

Gait is the *cyclic, coordinated rhythm of the body while moving on foot. Biometric gait recognition* is the verification, often referred to as authentication, and identification of an individual based on his or her walking style. Gait, as a biometric, authentication is the process of capturing the signal emanated from an individual's gait, determining whether it surpasses a threshold value for its similarity to known gaits stored in a database of previously observed signals, and confirming or denying the individual matches the individual's claimed identity. In order to build a database of gait signals and develop an effective classification technique, one must determine optimal gait processing techniques, discriminatory feature vectors, and efficient classifier settings.

A. BIOMETRICS

Biometrics are the unique biological and behavioral characteristics of an individual, distinguishable from those of other individuals. Originally described by Bianchi *et al.* [10], the different biomechanical characteristics of individuals, along with different kinematic strategies, that is, the individuals' control of energy oscillations in their bodies, allow gait to be reasonably categorized as a biometric. Following the terminology used in gait studies by Nickel [7] and Brandt [8], the term *genuine* will refer to an individual claiming an identity that matches his or her biometric sample. The term *imposter* will refer to a user with a biometric sample that does not match a claimed identity.

1. Biometric System

A biometric system is an automated system that captures, processes, and analyzes a biometric. The value from biometric systems comes from the capability to *verify* and *identify* an individual.

Identification of gait involves comparing an unknown gait sample to the entire database of known gaits. Identification is a one-to-N comparison for a database of N known gait samples. If the unknown sample, when compared to the database, surpasses a

threshold for similarity to a known sample, then the system should identify the unknown sample as the known individual.

Verification, also known as authentication, refers to the process of comparing an unknown gait sample to a gait sample that is claimed by the unknown individual. If the threshold is met for the similarity to the claimed sample, then the system should verify the individual's identity. As opposed to identification, verification is a one-to-one comparison.

a.   *Biometric System Performance*

For evaluating biometric performance, there exists a standard set of metrics for performance and evaluation; for an overview see the standards described by a report from the United States Military Academy (USMA) [11]. In this thesis, primary evaluation of classification performance will be derived from the *False Match Rate* (FMR) and *False Non-Match Rate* (FNMR) of each experiment.

The FMR is the proportion of zero-effort imposter attempts that are incorrectly classified as a match to the genuine subject. This metric helps describe the distinctiveness of a sample.

$$FMR = \frac{\text{Imposter Attempts Classified as Genuine}}{\text{Total Imposter Attempts}}$$

Conversely, the FNMR is defined as the proportion of genuine attempts falsely classified as imposter attempts. The FNMR can be used to assess the permanence of the biometric modality.

$$FNMR = \frac{\text{Genuine Attempts Classified as Imposter}}{\text{Total Genuine Attempts}}$$

In addition to the FMR and FNMR, the performance of a biometric system may be evaluated using the *False Acceptance Rate* (FAR) and *False Rejection Rate* (FRR). The difference between the FMR-FNMR and the FAR-FRR is that the latter takes the *Failure to Accept* (FTA) rate of the system into account. The FTA is the number of samples the system failed to successfully acquire due to problems with user presentation,

signal processing, feature extraction, or quality control. Since the data in this thesis is manually evaluated for quality prior to analysis, the FTA, and thus the FAR and FRR, will not be evaluated.

The *Equal Error Rate* (EER) will also not be reported in this thesis. The EER is the point where FAR and FRR are equal, or assuming FTA = 0, the point where FMR and FNMR are equal. Whereas this metric allows for a single value to report on the generalized performance of a biometric system, two systems with equal EER may have drastically different performance when comparing FMR-FNMR pairs under real-world operating settings and conditions, as described by Bromba [12]. For a security-centric authentication system, the primary goal is to minimize the FMR to prevent unauthorized access. Whereas minimizing FNMR is beneficial for system practicality, the goal is not to report only the point where FMR equals FNMR.

The performance of the techniques in this thesis will be evaluated and reported based on the *Total Error Rate* (TER). The TER is the sum of FMR and FNMR. Since the goal is to minimize the FMR, with an acceptable FNMR, this sum can be used to compare similarly low FMR results from different settings, while taking the usability of the system into account. Thus, a low TER will be evaluated as a well-performing feature vector.

B.   ACCELEROMETERS

The acceleration recording element, referred to as an *accelerometer*, is a small, embedded system found in almost every smartphone manufactured since 2005. The accelerometer can read data on three axes as shown in Figure 1: forward-backward (anteroposterior), up-and-down (vertically), and side-to-side (laterally). This provides raw acceleration data on X, Y, and Z axes. Depending on the device's orientation, gravity will affect one or more axes with a mean acceleration of approximately 9.81 m/s^2 toward the center of the Earth.
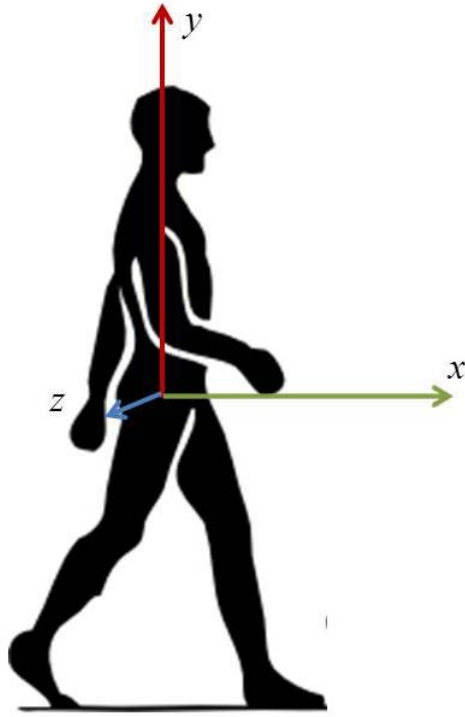
Figure 1. Human gait with anteroposterior (x), vertical (y), and lateral (z) directions.

C. GAIT RECOGNITION

The field of gait recognition involves the extraction of the unique characteristics of an individual's locomotion in order to map a gait signal to an individual. There are three primary methods for gait analysis as defined by Gafurov *et al.* [6]: Machine-Vision, Floor Sensors, and Wearable Sensors.

Machine Vision (MV) describes the use of optical sensors, such as cameras, on which computer vision techniques may be applied in order to detect and extract gait features. The benefit of MV gait recognition is that the subject does not have to explicitly interact with a device, nor even know a recording device is nearby. Previous works by Nixon *et al.* [13], Han *et al.* [14], and Liu *et al.* [15] have shown that optical systems, along with image feature extraction algorithms, may be applied to satisfactorily recognize individuals. This ability to identify individuals at a distance has implications in fields such as surveillance and security.

Floor Sensors (FS) involve the use of fixed ground sensors that can detect the characteristics of body mass movement over a fixed distance. Due to the overtness of a fixed FS, the techniques shown by Nakajima *et al.* [16] and Jenkins *et al.* [17] have applications in scenarios such as fixed facility access.

Wearable Sensors (WS) describe the method of attaching sensing devices to points on an individual's body in order to pull gait characteristics from the motion of body parts during locomotion. This thesis focuses on further developing the viable features of gait for a real-world approach to WS, specifically smartphone embedded accelerometers.

Wearable Sensors were chosen as the focus due to the ubiquity of smartphones, and their incorporated accelerometers, in our daily lives. Since the mid-2000s, smartphones have been manufactured with embedded accelerometers that are sensitive enough to detect the minute differences in an individual's walking rhythm. Additionally, these accelerometer sensors can be employed without a need for direct user involvement or change of normal daily activity. Thus, the employment of smartphone accelerometers for individual recognition has important implications on methods, including passive authentication, to secure smartphone data.

As early as Ailisto [18] in 2005, WS for gait recognition has been studied as a potential unobtrusive authentication method. Many previous studies collected gait data primarily from the hip position. One of the exceptions, Vildjiounaite [9] collected data from multiple positions; however, no attempt was made to combine the collected data from all positions in order to generalize a classifier for carrying-position-independent authentication. Table 1 provides a summary of the results of previous studies.

STUDY	POSTION	AXES	SEGMENT	FEATURES	RESULT
Gafurov [6]	Lower leg	X,Y,Z	None; Cycle-based	Histogram; Cycle length	EER=5%;9%
Nickel [7]	Hip	X,Y,Z, M	Time-based	BFCC, Min, Max, Std, Bin	TER=17.7%
Brandt [8]	Hip	X,Y	Time-based	BFCC	mTER=10.3%
Vildjiounaite [9]	Hand; Breast Pocket; Hip	X,Y,Z	Cycle-based	Correlation; FFT	EER=14.1% (Hip); 13.7% (Br)
Ailisto [18]	Lower back	X,Y	Cycle-based	Correlation	EER=6.4%; TER=12%
Mäntyjärvi [19]	Lower back	X,Y	Cycle-based	Correlation; FFT; Histogram	EER=7%; 10%; 19%
Gafurov [20]	Hip	X,Y,Z	Cycle-based	Euclidean Distance	EER=16%
Gafurov [21]	Right pocket	X,Y,Z	Cycle-based	Absolute Distance	EER=7.3%
Sprager [22]	Hip	X,Y	Cycle-based	Cumulant Coefficients	TER=7.4%
Rong [23]	Waist	X,Y,Z	Cycle-based	Dynamic Time Warping	EER=6.7%; TER=13.3%
Derawi [24]	Hip	X,Y,Z	Cycle-based	Cyclic-Rotation Metric	EER=5.7%
Gafurov [25]	Ankle	X,Y,Z	Cycle-based	Euclidean Distance	EER=1.5%
Holien [26]	Hip	X,Y,Z	Cycle-based	Dynamic Time Warping	EER=5.9%
Nickel [36]	Hip	X,Y,Z, M	Time-based	BFCC	EER=8.24%

Table 1.   Comparison of previous gait recognition research.

1.   **Gait Segmentation**

Since gait may be evaluated as a continuous signal, some form of segmentation must be performed in order to create discrete value for analysis and classification. The two primary methods are cycle-based segmentation and time-window segmentation. In

cycle-based segmentation, as used in Gafurov [6], Ailisto [18], Mäntyjärvi [19], and Sprager [22], the gait is assumed to be a periodic signal in which each gait cycle is the period after one foot touches the ground until that foot touches again. The first step in cycle-based segmentation involves identifying local minima and maxima over a designated period, thus requiring a peak-detection algorithm.

In Nickel [7] and Brandt [8], a time-window approach is introduced. Using a designated time-length window, the gait signal is segmented into windows of length *l*, with an overlap of *l*/2 for adjacent segments. With this approach, since gait is assumed to be periodic, each time segment is reasonably assumed to contain similar signal features. This approach requires fewer computational operations than cycle-detection and thus is more suited to use with mobile devices. Sliding window segmentation will be applied in this thesis.

2. Feature Extraction

Gait capture with WS involves the collection of a time series of raw accelerometer data points. As such, a feature vector consisting of each raw data point would be too large to make real-time processing realistic. However, a data reduction technique could be applied to the raw data points if they are evaluated as a discrete signal. Specifically, in Nickel [7] and Brandt [8], Mel-Frequency and Bark-Frequency Cepstral Coefficients are shown to be sufficiently unique descriptors of the signal characteristics.

The concept of the *cepstrum* was introduced by Bogert *et al.* [27] in 1963 as a heuristic technique for finding echo arrival times of composite signals, essentially defining the *cepstrum* as the spectrum of the log-spectrum of a function. Using this method, the cepstrum of a signal displays peaks where the original time waveform contained "echos" as described by Oppenheim *et al.* [28]. Since the *power cepstrum* described in Bogert [27] discards the phase information of the spectrum, Oppenheim [28] developed a *complex cepstrum.* The complex cepstrum, while still capable of echo detection, retains the phase information of the original wavelet, and may be used for wavelet recovery of the original signal. Thus, the coefficients from a discrete cosine

function applied to the *Mel* or *Bark* scaled cepstrum, can reduce the raw signal into a discrete and finite feature vector.

Mermelstein [29], in a preliminary experiment, showed that Mel-frequency Cepstral Coefficients (MFCC) could represent consonantal information in speech. Further, Davis [30] confirmed the selection of a compact frequency scale, such as *Mel,* which has linear frequency spacing below 1000 Hz and log-frequency spacing above 1000 Hz, can adequately represent speech with a small number of coefficients. The use of MFCC in speech and speaker recognition has since become standard. Both Nickel [7] and Brandt [8] showed the *Bark* scale, as opposed to the Mel scale, showed better results when applied to gait recognition.

D.   ACTIVITY DETECTION

The use of smartphone accelerometers to determine an individual's activity, be it sitting, standing, walking, running, or others, is a heavily studied area as presented by Kwapisz [31]. This field has application in many industries, including medical devices for fall detection, as studied by Zhang [32] and Dai [33]. While this thesis is focused on walking activity, some of the techniques used in activity detection will be applied.

The methods of determining an activity were shown by Tundo *et al.* [34] to perform better when the orientation of a device remains constant throughout the transition between activities. As such, many studies have required the attachment of the device in a known position on the human body. This, however, is not the normal behavior of individuals in the real world, who may place the device in a variety of pockets. Both Sprager [22] and Tundo [34] showed that with a baseline reading of the effect of gravity on the in-pocket accelerometer, a rotation can be applied to each subsequent accelerometer sample, thus re-orienting the device axes toward the strongest gravitational reading.

E.   DEVICE ROTATION

While experiments can attempt to control the orientation of a device used to capture gait data, in real-world application a device's base axes may be directed in an

unknown orientation due to variances in the size and shape of a carrying location, such as a pocket. If a device is accidentally rotated, with respect to gravity, it was shown in Tundo [34] that normalizing all of the raw data mathematically by rotating from the device's reference frame to a gravity-based reference frame, as shown in Figure 2, can provide more accurate classification results for activity detection.
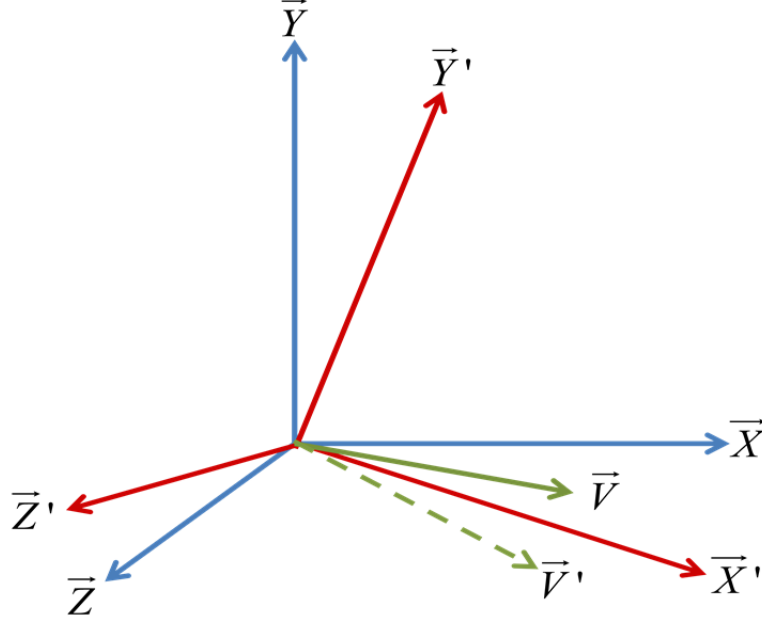
Figure 2.   A device's reference frame (blue) rotated toward gravity (red).

If a vector $\vec{v}$ exists in the device's reference frame F, it can be transformed to $\vec{v}'$ in the gravity reference frame F' by multiplying it by a rotation matrix R, representing the transform from F to F'. Sprager [22] employed a calibration technique that captured each subjects' stationary data with the collection device in position for a period prior to the start of the walk. This initial stationary period was averaged to calculate a rotation matrix based on gravity, which was then applied to each subsequent sample during the walk. The effect of a similar calibration technique will be evaluated in this thesis.

F.   MACHINE LEARNING

Machine learning is the method of giving a computer the ability to *train* to better perform a task without requiring explicit programming. Learning may be performed

13

either supervised or unsupervised. In supervised learning, the machine learns to make predictions based on previously seen events. Supervised learning has the benefit that, once properly trained, the machine may begin providing accurate predictions immediately after initial training. In unsupervised learning, the machine does not have knowledge of previous events, but instead must possess a method of receiving feedback on the accuracy of its predictions. With proper feedback, the unsupervised method can further improve its accuracy through exposure to more events. While unsupervised learning may suffer from a "learning curve" in early predictions, it benefits from not requiring an extensive database of prior knowledge.

Since we are concerned with authenticating a known individual, this thesis will focus solely on supervised learning. The machines employed will be trained on known instances of an individual's gait in order to classify future instances from this prior knowledge.

1.   Machine Learning Techniques

Development and employment of optimal algorithms to perform machine learning tasks is a heavily studied field. Here, we evaluate and compare the classification accuracy of two common machine learning algorithms, Support Vector Machine (SVM) and k-nearest neighbor (kNN). Both techniques require two sets of data: training and testing sets. Both sets contain multiple vectors, each representing a *case*. Each case contains a group of attributes of that case known as *features.* Each case is also given a *label* of what class the case is a member. The goal of each technique is to build a model from the training set that most accurately predicts the labels of the cases in the testing set.

a.   *Support Vector Machine*

Support Vector Machine (SVM) was chosen for this thesis as it has shown good performance in previous gait recognition research, including both Nickel [7] and Brandt [8], but also because this algorithm works particularly well for binary classification tasks, or the process of describing an event as one of two known classes.

SVM works by determining a hyperplane of n dimensional space, with n equal to the size of the feature vector, which best separates two or more classes. It does this by determining the maximum margin between the hyperplane that best divides the classes, and the *support vectors,* or the closest data points to the hyperplane. Using $\vec{x}$ as the point (vector) and $\vec{w}$ as the weights, then the hyperplane may be defined as $\vec{w}^T \vec{x}_k + b > 0$ for all $\vec{x}_k$ of one class and $\vec{w}^T \vec{x}_j + b < 0$ for all $\vec{x}_j$ of the other class as explained by Lewis [35]. If each class is labeled as $y_k \in \{0,1\}$ with 1 being a positive case and 0 a negative case, then the equation for the hyperplane is $y_k(\vec{w}^T \vec{x}_k + b) \geq 0$.

Since (x_k, y_k) is known for all training cases, then this equation may be used to solve for $\vec{w}, b$ which gives the hyperplane. In cases where a clear division between the classes does not exist, the SVM employs a *slack variable* which provides an allowance of data points to lie on the wrong side of the division. Thus, the goal of the SVM is to minimize the slack variable and maximize the margin between support vectors. The SVM will either be able to draw a clear divide between the classes with no error, or it will have some error with data points on the wrong side of the divide. In either case, the same equation may be used to determine the maximum margin between data sets. If the data can be clearly separated, then the error penalty $C = 0$. If there is some error, then $C > 0$.

$$\min_{w,b,\xi} \frac{1}{2} \|w\|^2 + C\sum_i^n \xi$$

Here ξ is the slack variable and the entire term $C\sum_i^n \xi$ is the soft margin for the SVM. In order to determine the optimal C and ξ, a logarithmic grid search may be used, which evaluates the performance of the SVM on the training data using all possible pairs of $C = \{2^{-5}, 2^{-3}, 2^{-1}, 2^1, 2^3, 2^5\}$ and $\xi = \{2^{-15}, 2^{-10}, 2^{-5}, 2^0, 2^5\}$, then scores the outputs to determine the optimal. Grid search is a computational intensive task, as the training data is first divided into equal-sized segments, then N SVMs are built with N-1 training sets and one test set, for each of the N segments.

b.   ***k-Nearest Neighbor***

K-nearest neighbor (kNN) was chosen since in at least one previous study by Nickel [36], kNN showed better performance in accurately classifying gait than SVM. The kNN technique involves calculating the distance between a test case, which is a vector of attributes, and stored training cases. In this thesis, as with Nickel [36], the Euclidean distance will be used. The k-nearest neighbors, by distance, of the test vector then "vote" based on their labels, and the majority label of the k-neighbors is applied to the test vector. If there is an even number of positive and negative neighbors, then the genuine label, by default, is applied to the test case.

2.   Machine Learning Tools

Orange [37] is a comprehensive, open-source toolbox for machine learning and data mining. Based on Python, it provides the user the ability to quickly write scripts to perform a multitude of tasks, including data management, classifier construction, calibration, prediction, evaluation, and visualization, using built-in functions. The data management and preprocessing functions allow loading data, sampling data, filtering, scaling, attribute selection, set construction, and saving data. In classifier construction, there are functions for training and testing SVMs (based on LibSVM), kNN, Decision trees, and many others. The prediction, evaluation, and visualization functions use trained classifiers to predict classifications of a training set and score the output, which may then be displayed graphically for the user. Due to the scale and flexibility of the Orange toolbox, the classification tasks in the thesis are performed using a custom Python script leveraging Orange's functionality.

III. METHODOLOGY

This chapter discusses the software design considerations and describes the equipment used for the experiments. First, the method developed to extract accelerometer data from the Android device is explained. Then, a description of the signal processing and feature extraction procedure is provided. Finally, the implementation of the gait classifier is discussed.

A. DATA COLLECTION

The gait database for these experiments consists of raw accelerometer data collected from a LG Nexus 4. The Nexus 4 comes embedded with an Invensense MPU-6050 Six-Axis gyroscope and accelerometer, which measures accelerometer data in three directions as illustrated in Figure 3.
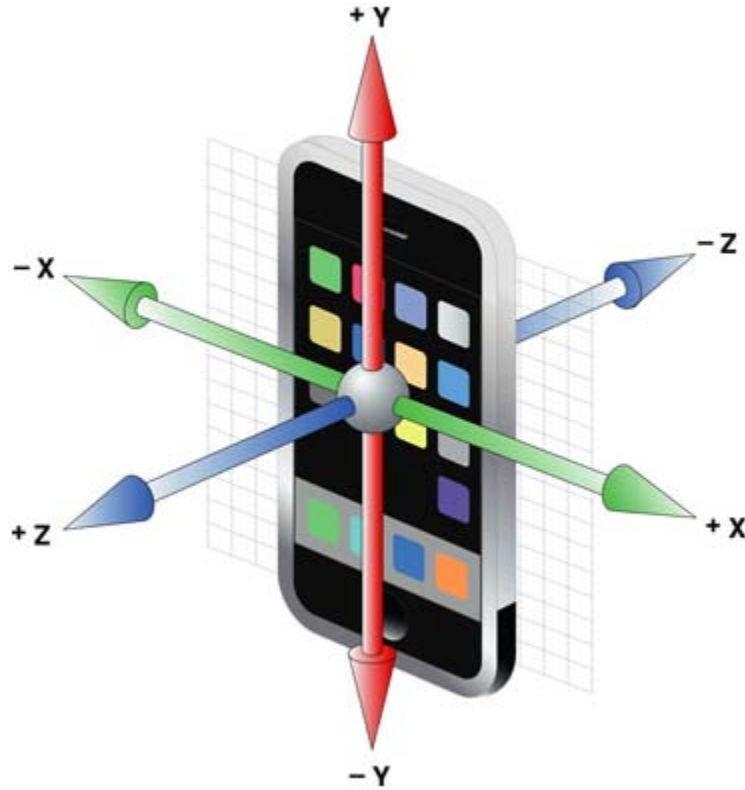
Figure 3.   The accelerometer axes in the device's frame of reference, from [38].

An application was written using the Android SDK to extract raw accelerometer data using the Android SDK's *onSensorChanged* method. As the accelerometer reading changes over time, the data is written to an internal SQLite database containing an instance identifier, a time-stamp, and the accelerometer magnitude in each of the three directions.

To build a database of representative, real-world gait accelerometer measurements, 23 subjects were invited to participate in data collection. Summaries of the group demographics appear in Table 2 and Table 3. Note that the majority of participants were healthy males between the ages of 26 and 34 with an average height of 5′10″, minimizing the effects of age, gender, and health on classification.

AGE	MALE	FEMALE
26–28	8	0
29–31	5	2
32–34	6	1
> 35	1	0

Table 2.   Summary of subject ages.

HEIGHT	MALE	FEMALE
< 5'6"	0	1
5'6"-5'9"	5	2
5'9"-5'11"	11	0
6'-6'2"	4	0

Table 3.   Summary of subject heights.

A data collection *session* was conducted with the device in each of three locations: back pocket, front pocket, and hip holster. All carrying positions were on the right side of the body and each subject wore business casual clothing and shoes. In each

position, the device was carried with the screen facing toward the body and the top of the device toward the ground. This ensures the raw orientation of each axis is consistent, in relation to positive and negative values, no matter the carrying position. This is necessary for rotation normalization during signal processing.

After loading the data collection application, each subject placed the device first in the back pocket. For approximately eight seconds, the subject was asked to remain still in order to collect gravity calibration data. At the end of the approximately eight seconds, the subject was tasked to walk at a comfortable pace for approximately 20 seconds on a straight path. Following this approximately 30 seconds of total data collection, the database was saved, the device reset, and the experiment conducted again with the device in the front pocket and finally in the hip carrying position.

In order to ensure enough data was available for training and testing sets, following the completion of the first round of data collection walks, each walk was conducted a second time for a total of six, 30-second data collection sessions for each subject providing a total of approximately 4200 seconds of raw data.

B.   SIGNAL PROCESSING

Following the completion of the data collection, the raw database is visually inspected using database plotting software. After visual verification of the X, Y, and Z magnitudes over time, the database can be loaded into a custom Python script, leveraging *Scipy and Numpy* [39] libraries for the following manipulations.

1.   Interpolation

Due to a limitation in the Android API, the only time the data is pulled from the accelerometer is when the sensor changes. As this change may not occur at a fixed interval, and since higher priority processes may interrupt the polling of the accelerometer, the raw data is not stored at a uniform sampling rate. Table 4 shows a sample of the raw data from the calibration portion of a sample gait collection session. Observe the inconsistent time deltas, in nanoseconds, between data points.

TIME DELTA	x	y	z
20141600	-0.358291625976562	-9.813888549804687	2.42636108398437
20141602	-0.364242553710937	-9.876968383789062	2.41921997070312
20141602	-0.352340698242187	-9.779373168945312	2.39541625976562
20141600	-0.391616821289062	-9.769851684570312	2.44540405273437
20172120	-0.398757934570312	-9.805557250976562	2.47991943359375
20172120	-0.366622924804687	-9.788894653320312	2.42755126953125
20141600	-0.344009399414062	-9.817459106445312	2.44778442382812
20080566	-0.367813110351562	-9.853164672851562	2.391845703125
20141602	-0.322586059570312	-9.847213745117187	2.41207885742187
20141600	-0.321395874023437	-9.811508178710937	2.44302368164062

Table 4.   Sample of raw gait data with time delta between samples.

In order to ensure we are comparing controlled data sets, interpolation to a fixed rate is necessary. As each subject's data was collected for approximately 30 seconds, we interpolate to 1500 samples in the 30 second sessions to extract data at 50 Hz. In Brandt [8], 50 Hz was shown as an adequate rate for gait discrimination. Raw data points beyond 30 seconds are dropped in order to ensure the same number of samples for each walk session. This leaves us with the blue signal, as depicted in Figure 4.
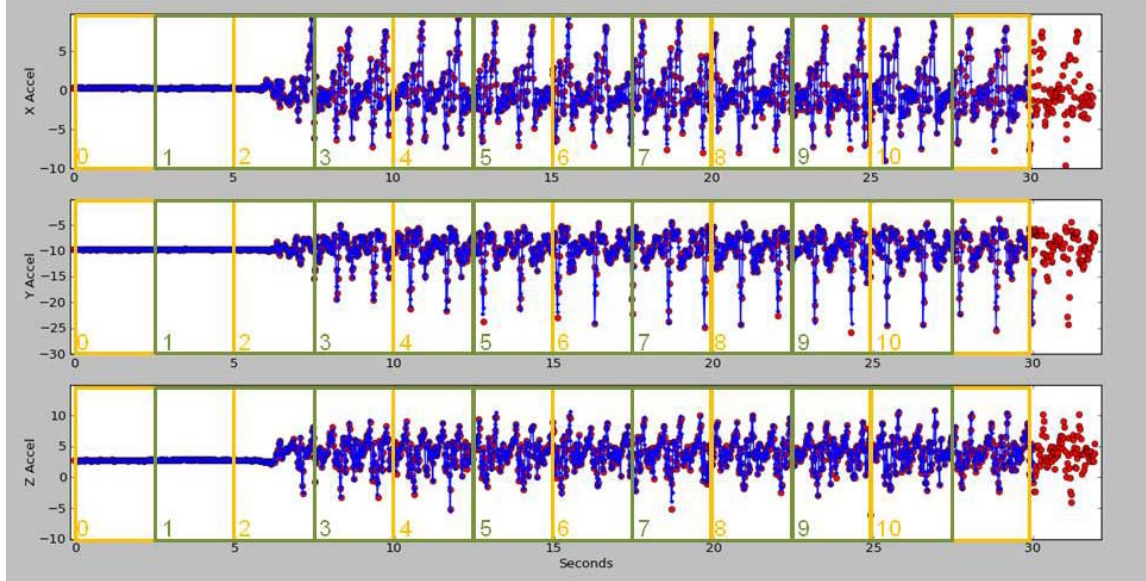
Figure 4. The raw data points (red) and interpolated signal (blue) in X, Y, and Z axes with the overlapping segment boundaries overlaid in gold and green.

2. Segmentation

With each session of raw data interpolated to 1500 samples, each session is divided into discrete segments. In performing non-cyclic gait classification, Nickel [7] and Brandt [8] showed equal length segments with 50% overlap provide low FMR and FNMR. The raw signal of length l is split into segments of time t with a distance d between the start of consecutive segments. Since our data collection and experimental method is similar to that in Brandt [8], we use their optimal segment lengths where $t = 5$ seconds, which at 50 Hz is 250 samples per segment, and $d = 2.5$ seconds.

Starting with the approximately 30 seconds of initial raw data, each session is thus divided into 11 segments. Due to the approximately eight-second resting calibration period of each session, the first three segments are assumed to be non-walking segments and discarded after normalization is complete, leaving eight segments per subject for the classification task.

3. Normalization

Note the presence of gravity on the Y-axis in Figure 4. This noise on the raw signal may potentially interfere with the recognition of an individual's gait and instead

21

classify a user based on the way the device is carried. Due to this potential, classification will be performed on gaits normalized using several techniques.

As a baseline, experiments are first run to classify gait without normalizing the raw data. Following this, the normalization of Nickel [7] and Brandt [8], which was used to remove accelerometer noise and allow an evaluation of a zero-crossing metric, is verified by averaging the signal in each segment to yield μ. This average is subtracted from each data point to get a zero-normalized value at time t.

$$s_k'(t) = s_k(t) - \mu_k \text{ for } k = \{x, y, z\}$$

Additionally, the effect of normalizing the raw data by an axis-rotation, as proposed by Tundo [34] and Cooke [40], is evaluated. The direction and magnitude of the axis rotation can be determined from the mean magnitude of the acceleration of each axis during the eight second "at-rest" calibration potion of the data collection session. After calculating the mean, a direction and magnitude of axis-rotation is determined and applied to each data point. While a rotation matrix can perform this task, for this thesis a quaternion rotation is performed due to calculation speed and efficiency.

Finally, after evaluating the effect of the rotation on classifier results, the rotation technique is applied, followed by a scaling to zero. The average for each segment is calculated for each rotated segment, and this mean subtracted from each data point to center all data points near zero.

C.   QUATERNION ROTATION

Contrary to using a 3-by-3 matrix to represent rotation in three dimensions, a four-dimensional quaternion vector may be used instead. Quaternions offer several advantages over using a rotation matrix, including compact representation and storage requirements. Further discussion of the advantages of employing quaternions for rotations can be found in Dam *et al.* [41]. Instead of applying many elementary arithmetic operations on a 9-element matrix, we instead operate on the 4-element quaternion represented by the following linear combination.

$$q = q_0 + q_1 i + q_2 j + q_3 k, \text{ where } i^2 = j^2 = k^2 = ijk = -1$$

22

The process of rotating an angle by a quaternion involves building a quaternion rotation matrix. First, the axis-vector must be produced from the cross-product of the initial gravitation-vector and the desired gravity vector.

$$A = \vec{v_i} \times \vec{v_f}$$

Since the initial vector is $\vec{v_i} = Xi + Yj + Zk$ and the desired gravity vector is $\vec{v_f} = X'i + Y'j + Z'k$, then the dot product may be used to derive the angle of rotation.

$$\vec{v_i} \cdot \vec{v_f} = \left\| \vec{v_i} \right\| \left\| \vec{v_f} \right\| \cos(\alpha)$$

$$\vec{v_i} \cdot \vec{v_f} = XX' + YY' + ZZ'$$

The desired gravity vector is equal to (0.0, -9.81, 0.0) so X' and $Z' = 0$ and $\left\| \vec{v_f} \right\| = Y'$, therefore the angle of rotation is:

$$\alpha = \arccos\left(\frac{Y}{\left\| \vec{v_i} \right\|} \right).$$

The angle of rotation and the axis-vector elements can now be used to build the quaternion rotation matrix using the quaternion rotation equations described in Cooke [40].

$$q_0 = \cos\left(\frac{\alpha}{2} \right)$$

$$q_1 = \sin\left(\frac{\alpha}{2} \right) A_x$$

$$q_2 = \sin\left(\frac{\alpha}{2} \right) A_y$$

$$q_3 = \sin\left(\frac{\alpha}{2} \right) A_z$$

23

$$R(q_0, q_1, q_2, q_3) = R = \begin{pmatrix} 1 - 2(q_2^2 + q_3^2) & 2(q_1 q_2 - q_0 q_3) & 2(q_0 q_2 + q_1 q_3) \\ 2(q_1 q_2 + q_0 q_3) & 1 - 2(q_1^2 + q_3^2) & 2(q_2 q_3 - q_0 q_1) \\ 2(q_1 q_3 - q_0 q_2) & 2(q_0 q_1 + q_2 q_3) & 1 - 2(q_1^2 + q_2^2) \end{pmatrix}$$

Considering an initial vector represented by $[X, Y, Z]^T$, multiplying this vector with the rotation matrix produces the rotated vector with respect to the desired gravitation vector.

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = R \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

D.   FEATURE EXTRACTION

Both Nickel [7] and Brandt [8] showed that Mel and Bark Frequency Cepstral Coefficients (BFCC), commonly exploited in speech recognition tasks, perform better for gait classification than statistic features. Thus, instead of using common statistical features such as max, min, mean, and standard deviation, this thesis will instead use the BFCC calculated using the optimal parameters as described in Brandt [8]. Bark scale was first described by Zwicker [42] in 1961. A detailed discussion of MFCC and BFCC construction can be found in Rabiner [43]. An overview of the BFCC process is shown in Figure 5.
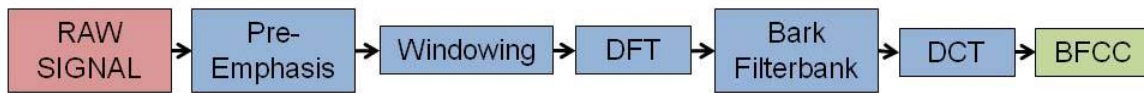
Figure 5.   The BFCC process.

1.   Pre-emphasis

The first phase is pre-emphasis of the raw signal. In this step, higher frequencies are emphasized by increasing the energy of the signal in the higher frequency bands. The equation for this indicates that the pre-emphasized sample is equal to the raw sample minus 97% of the previous sample.

$$A'[n] = A[n] - 0.97 * A[n-1]$$

The raw input signal is constantly changing over the length of the segment; however, in order to simplify calculations, it can be assumed the signal does not change significantly over a shorter window.

2. Windowing

Note the term *segment* is used to describe a portion of the original signal and *window* is used to describe a portion of the segment. Thus, the 250 sample segment is further windowed, again with 50% overlap.

3. Discrete Fourier Transform

Since each window is reduced enough to be easily described by a few coefficients, a Fourier Transform, in the form of a FFT for speed of computation, is performed on each window.

4. Bark Filterbank

In order to smooth the spectrum and emphasize the meaningful frequencies, the spectral components are divided into frequency bins according to the Bark scale. The Bark scale, like the Mel scale, is based on findings that, in speech, lower frequencies are perceptually more important than higher ones. Thus, the Bark filterbank is applied to the frequency outputs of each FFT using the following conversion formula described by Traunmüller [44].

$$Bark(f) = \frac{26.81f}{1960 + f} - 0.53, \text{ where } f \text{ is the vector of frequencies from the FFT}$$

This reduces the calculated FFT spectrum into a reduced set of energy values. The optimal results, as produced by Brandt [8], reduce the FFT spectrum to 40 values, of which the log is taken.

5.   Discrete Cosine Transform

Finally, a Discrete Cosine Transform (DCT) is applied to each of the 40 log-energies. Since the Bark-frequency vectors calculated for each window are highly correlated, a DCT is used as an approximation of the Karhunen-Loeve transform, which decorrelates the vectors and thus reduces the number of parameters in the system Logan [45]. The DCT yields 40 coefficients describing the original signal. Due to the spacing of the frequencies in the filterbank, the coefficients past 13 contain little information beyond noise and are thus discarded. With 13 coefficients for each of the windows, the mean value for each coefficient is returned as a vector of the BFCC for the input segment.

Dan Ellis' [46] MFCC implementation for MATLAB, as used in Nickel [7] and Brandt [8], was chosen for this thesis. This implementation calculates cepstral coefficients as described and includes an option to use the Bark-scale filterbank vice the Mel-scale. Table 5 displays the parameter settings.

PARAMETER	VALUE
Window Length	0.007
Window Hop Time	0.0002
Sampling Rate	16000
Minimum Frequency	0
Maximum Frequency	1200
Pre-emphasis Filters	0.97
Number of Spectral Bands	40
Number of Cepstral Features	13
Cepstral Liftering	None
Cepstral Scale	Bark

Table 5.   Dan Ellis' script settings for BFCC calculation

E.   SIGNAL CLASSIFICATION

In order to determine the best performing combination of the features being investigated, a binary classifier is developed. The purpose of the classifier is to separate instances of a genuine user from that of the many imposters: thus, it is a one-to-many classifier. Since the goal is to determine the best performing features in any case, for each experiment N one-to-many classifiers are built for each of N subjects. The results on each classifier are combined to get an average TER for the feature set.

The data is first captured and processed as described. Once the features are extracted for all samples, half of the resulting feature vectors are used for training and the other half are used for testing. In all experiments, the training occurs on the samples from the first walking sessions and testing is performed on the second walking session. No samples from the training set appear in the testing set.

In order to determine the best performing classifier for the test, each experiment is run concurrently on both a SVM and kNN trained and tested on the same data sets. Both techniques are implemented in a Python script leveraging the Orange libraries. Prior to loading the feature vectors, the vectors are scaled to values between zero and one in order to reduce the influence of higher value attributes on others. In order to ensure scaling by the same amount on the training and testing sets, allowing values in each set to vary by different ranges, the scaling amount from the training set must be saved and applied to the data in the testing set regardless of its internal range variance. This method is presented as a best practice by Hsu [47].

Additionally, unless noted otherwise, all classification tasks are performed on homogenous carrying positions: that is, the classifier will be trained on walks from the same carrying positions as those in the test set. This is done in order to evaluate the signal classification effect from different positions and to determine which classifier, normalization technique, and primary axis are most discriminatory.

1.   Classifier Settings

The first step in building the Support Vector classifier requires the determination of the optimal kernel parameter γ and the penalty parameter C. The Gaussian radial basis function (RBF) was chosen as it was shown by Brandt [8] and Hsu [47] as well performing for gait classification tasks. The SVM implementation in Orange uses the automatic parameter selection function from LibSVM, performing a grid search on all pairs of logarithmically-spaced γ and C values during cross-validation in order to determine the optimal pair.

In building a kNN classifier, the two parameters that most affect performance are the distance function and the number of neighbors, k. In keeping with Nickel [36], the Euclidean distance is used and $k = 8$ was selected as each walking session is previously determined to contain eight segments.

2.   Experimental Method

In order to evaluate the effect of carrying position and normalization technique, multiple combinations of feature selection mixtures were evaluated. For each carrying position (hip, back pocket, and front pocket), and for each axis (x, y, and z), each of our normalization techniques (no normalization, zero-scaling, rotation, and zero-scaled rotation) was evaluated. Additionally, for each position-normalization technique, the X and Y axes and all three axes are combined separately, as these combined features were shown to perform well in previous studies (see Table 1). The important information when combining axes is the overall effect of the combined magnitude of the individual axes. In order to calculate this combined magnitude, at each time t the data points from each axis of interest ($s_x(t)$, $s_y(t)$, and $s_z(t)$) is treated as a vector with the magnitude of its value in the direction of the respective axis. The Euclidean norm of these axis vectors is calculated for each t and the features extracted from the combined segment vector s_{comb}.

$$s_{comb}(t) = \sqrt{s_k(t)^2} \text{ for } k = \{x, y, z\}^+$$

Following the structure of Brandt [8], the initial baseline was developed using and SVM on data originating at the hip-carrying position, processed with the zero-scaling normalization technique for each individual axis. The TER results are showing in Table 6.

POSITION	AXIS	FMR	FNMR	TER
HIP (H)	X	0.0217	0.6250	0.6467
	Y	0.0326	0.4239	0.4565
	Z	0.0440	0.4837	0.5277

Table 6.   Baseline results for individual axes.

It can be seen that the individual Y and Z axes are the best performing. When combining axes, the X and Y axes show good performance (see Table 7), which is similar to Brandt [8], however, not an improvement to the individual Y-axis.

POSITION	AXIS	FMR	FNMR	TER
HIP	XY	0.0292	0.4674	0.4966
	XYZ	0.0341	0.4837	0.5178

Table 7.   Baseline results for combined axes.

3.   Voting Scheme

Though the baseline results are computed for each segment of an individual's walk session, the goal is to authenticate a user over a period of time. In the current settings, with a FNMR near 50%, the genuine user is rejected about half the time. For a usable system, this result is unacceptable. In order to improve the usability of the classifier a voting scheme is implemented similar to the ones described by Nickel [7] and Brandt [8].

Instead of authenticating on each individual segment, several consecutive segments V are used to authenticate a period of time so long as a goal number G of the segments are classified as genuine. For the following experiments, since the number of genuine segments is known to be eight, in order to ensure only genuine or only impostor segments are included in the authentication period, the V must be a whole number divisor of eight. In order to set the goal number, the baseline classification settings are used to empirically evaluate the best performing V and G numbers to be used for following experiments. The abridged results of the voting scheme are shown in Table 8, where it can be seen that XY is the best performing axis. While the FMR increased by six percentage points, the TER decreased by over 50%.

AXIS	METHOD	V	G	FMR	FNMR	TER
X	SVM	8	1	0.0652	0.3478	0.4130
	kNN	8	1	0.1759	0.1304	0.3063
Y	SVM	8	2	0.0613	0.1739	0.2352
	kNN	8	3	0.0692	0.1739	0.2431
Z	SVM	8	2	0.0711	0.3043	0.3754
	kNN	4	1	0.1314	0.2826	0.4140
XY	SVM	8	1	0.0988	0.1087	0.2075
	kNN	8	2	0.0850	0.1087	0.1937
XYZ	SVM	8	1	0.0929	0.1739	0.2668
	kNN	8	2	0.0830	0.1739	0.2569

Table 8.   Performance of voting scheme on baseline parameters.

Table 9, showing the mean TER for the combined performance of both SVM and kNN on all axis combinations, indicates that the use of eight votes, with a goal of two genuine, has the best overall performance and is used in further experiments. Appendix A includes the results of all voting optimization experiments.

30

V	G	mTER
1	1	0.4981
2	1	0.3945
2	2	0.6025
4	1	0.3352
4	2	0.3989
4	3	0.5206
4	4	0.7394
8	1	0.3132
8	2	0.3118
8	3	0.3336
8	4	0.3838
8	5	0.5011
8	6	0.5925
8	7	0.6951
8	8	0.8571

Table 9.   Mean TER of SVM and kNN performance with different voting parameters.

Since each session in the data set includes eight segments of five seconds each, several cases must be addressed. For example, if only a single segment in a series is classified as genuine, should the entire period be recognized as authentic? If so, does it matter where this genuine segment falls in the walking period be it beginning, middle, or end? Thus, cases may represent times of positive authentication, such as when a genuine user walks and stops, or walks and takes the device out of the pocket, or negative authentication such as when a device is taken from the genuine user in the middle of a walking period. The study of these settings is particularly interesting and we leave their exploration for future research. Following Nickel [7] and Brandt [8], we evaluated the bundling of segments and determined the entire 20 second walk from bundling 8-segments is a practical unit of classification for our data set.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. ANALYSIS AND RESULTS

In the process of developing a gait authentication system and analyzing feature processing methods, several important results were determined. First, as described previously, it was found that the implementation of a voting scheme can improve the usability of a gait authentication scheme by reducing the system's FNMR with acceptable FMR trade-off. For this data set, requiring two positive votes out of eight, five-second segments showed the best performance improvement. These settings were used for evaluation and comparison for the rest of the experiments.

A. CLASSIFIER EVALUATION

In determining the best performing classifier technique and signal processing methods, 60 experiments were conducted involving 20 different *mixtures* of feature vectors and processing techniques, as described by Table 10. Each mixture's performance was evaluated for each of the three carrying positions.

AXIS	NORMALIZATION
X	None
	Zero-Scaling
	Rotation
	Rotation and Zero-Scaling
Y	None
	Zero-Scaling
	Rotation
	Rotation and Zero-Scaling
Z	None
	Zero-Scaling
	Rotation
	Rotation and Zero-Scaling
XY	None
	Zero-Scaling
	Rotation
	Rotation and Zero-Scaling
XYZ	None
	Zero-Scaling
	Rotation
	Rotation and Zero-Scaling

Table 10.   Experimental mixtures for each carrying position.

While all experiments were run on both SVM and kNN classifiers, in 51 of 60 experiments kNN outperform SVM. Table 11 shows the mean TER of the classifiers on all experiments. This concurs with the results of Nickel *et al.* [36] and suggests that, in general, kNN is a better performing technique for gait classification. Full results of all the experiments are included in Appendix B with graphical representations of per position performance in Appendix C.

	SVM	kNN
mTER	0.3913	0.3010

Table 11.   Means of classifier performance on all experiments.

Using kNN on all experimental mixtures, we developed a scatter plot of classifier performance that provided us a method to visually determine viable mixtures (see Figure 6). As the intent is to evaluate the performance regardless of device position, and observing that the inter-position variance is relatively low, with a mean variance of 0.0068, we perform further analysis on the mean TER of all positions (see Figure 7).
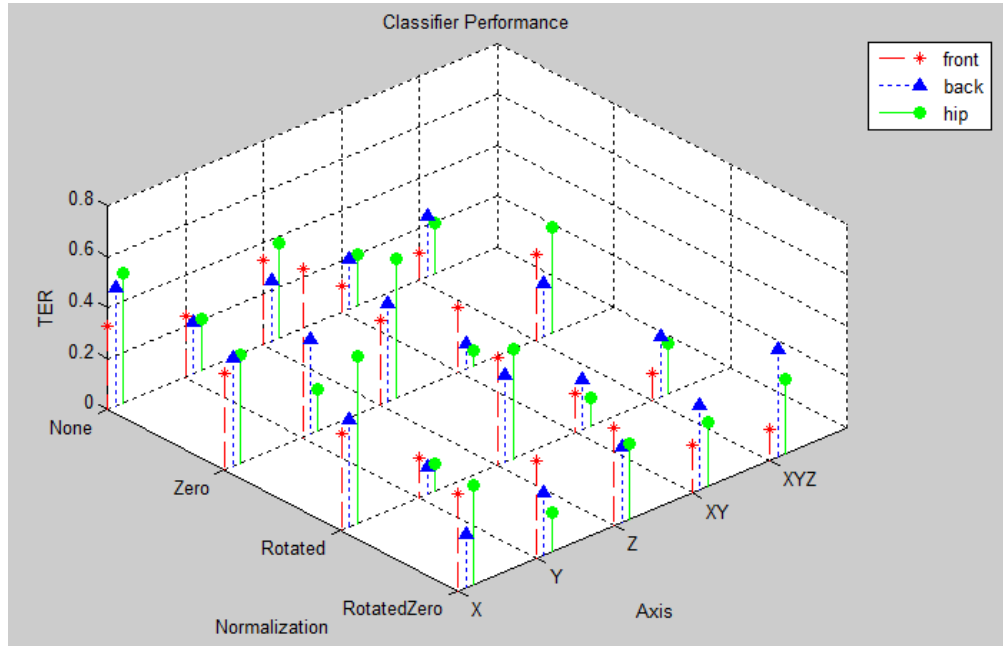
Figure 6.   3D scatter plot depicting the performance of kNN, by TER, on all experimental mixtures.
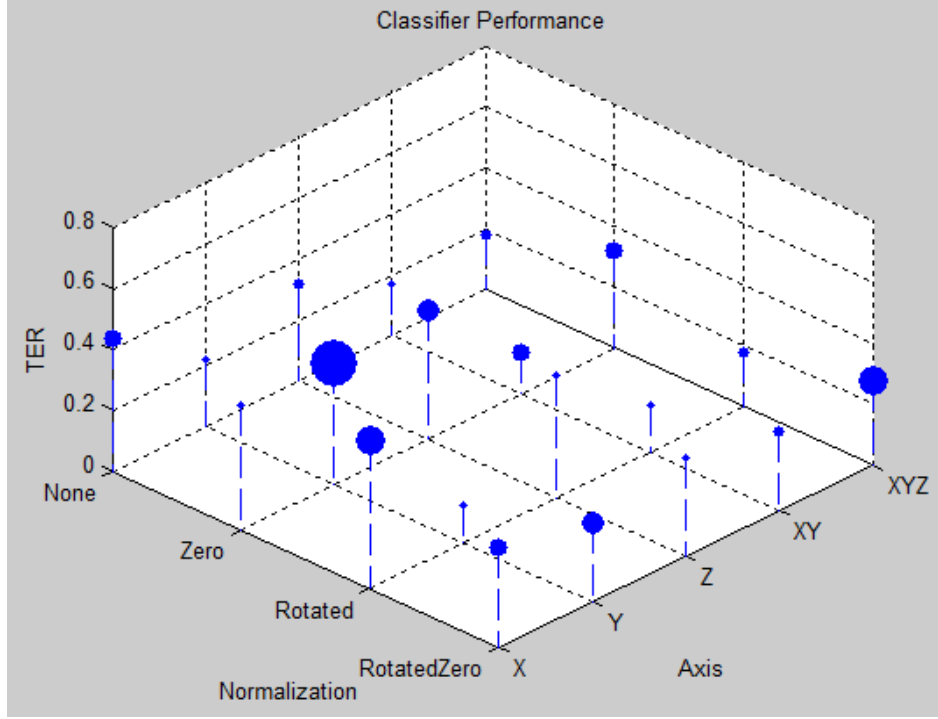
Figure 7.   3D scatter plot of the inter-position mTER where large points represent
a large inter-position variance (see Appendix D).

1.   Normalization Analysis

To evaluate the performance of the normalization techniques on each axis, Figure 7 allowed us to observe data across the intersections of normalization and axis settings. We first observed that the best performing data point, by lowest mTER, occurred at 12.6% on the Y-axis when normalized by a gravity rotation. The XY, and XYZ axes have similar performance, 15.4% and 18.0%, respectively, indicating a dependence on the Y-axis.

The zero-scaling normalization technique displayed the worst inter-position variance regardless of axis at 0.0129, though the second-best mTER of 13.8% occurs on the zero-scaled XY-axis. Additionally, the single best position-dependent data point occurred on the zero-scaled XY-axis, from the hip carrying position, at 6.13% TER. Of note, this mixture was the primary one used by Brandt [8].

When no data normalization was performed, the data showed the lowest inter-position variance at 0.0029 and an average TER across axes centered at 26.3%. Though

the rotated and zero-scaled data showed a poor inter-pocket variance at 0.0072, it has the lowest inter-axis variance with all data points clustered around 28% TER. Results are summarized in Table 12.

	CENTER	VARIANCE
None	26.3%	0.0102
Zero-Scaled	33.8%	0.0113
Rotated	26.9%	0.0212
Rotated and Zero	28.9%	0.0009

Table 12.   Normalization cluster centers and inter-axis variance across axes.

2.   Axis Analysis

Similar to the analysis of the normalization techniques, we used Figure 7 to make observations of the effect of axis-selection on classification. We first observed that the Y-axis had the lowest mTER of the individual axes, but the highest variance across normalization techniques. The combined XY-axis had the lowest mTER at 17.9% and the lowest variance across normalizations. Again, this agreed with Brandt's [8] selection of the XY-axis for gait authentication. If we looked at only zero-scaled data, the XY-axis also showed the best mTER of 13.7% though it had the highest inter-position variance, 0.0069, of XY-axis data across normalization techniques. The XY-axis data with the lowest inter-position variance, 0.0013, occurred on gravity rotated data achieving an mTER of 15.4%. This indicates position-independent gait authentication may perform more consistently on XY-axis data normalized by a gravity rotation.

	CENTER	VARIANCE
X	41.5%	0.0035
Y	25.0%	0.0097
Z	36.5%	0.0021
XY	17.9%	0.0021
XYZ	23.9%	0.0037

Table 13.   Axis cluster centers and inter-normalization variance
across normalization techniques.

B.   CARRYING POSITION INDEPENDENT CLASSIFICATION

Samples received from different carrying positions were not combined in any form, as receiving signal from multiple positions would be impossible with a single device. Instead, in order to determine the most realistic method of authenticating gait data from different carrying positions, analysis was performed using the previously determined best performing axis (XY) and two best normalization techniques (zero-scaling and gravity rotation) and mixing training and testing sets from the different carrying positions. Based on our results indicating a strong dependence on the Y-axis and our best performance agreeing with the selection of the XY axis by Brandt [8], we did not evaluate the performance of the following experiment using other combinations of axes.

First, the classifier was trained and tested on data from the individual carrying positions. Then, the classifier was trained on data from all the carrying positions, and testing was performed on data from the individual carrying positions. We present the results in Figure 8 and Figure 9, using the zero-scaling and gravity rotation, respectively.

It is apparent that classifiers trained and tested on data originating from the same carrying position show the best performance. The effect of the normalization techniques is particularly interesting. Though the performance when comparing data from the same positions is slightly worst using the gravity rotation technique, the performance of position combination techniques all improve. The performance of training on all data and

testing on a single position is equivalent, or better, than training and testing on the back pocket position alone.
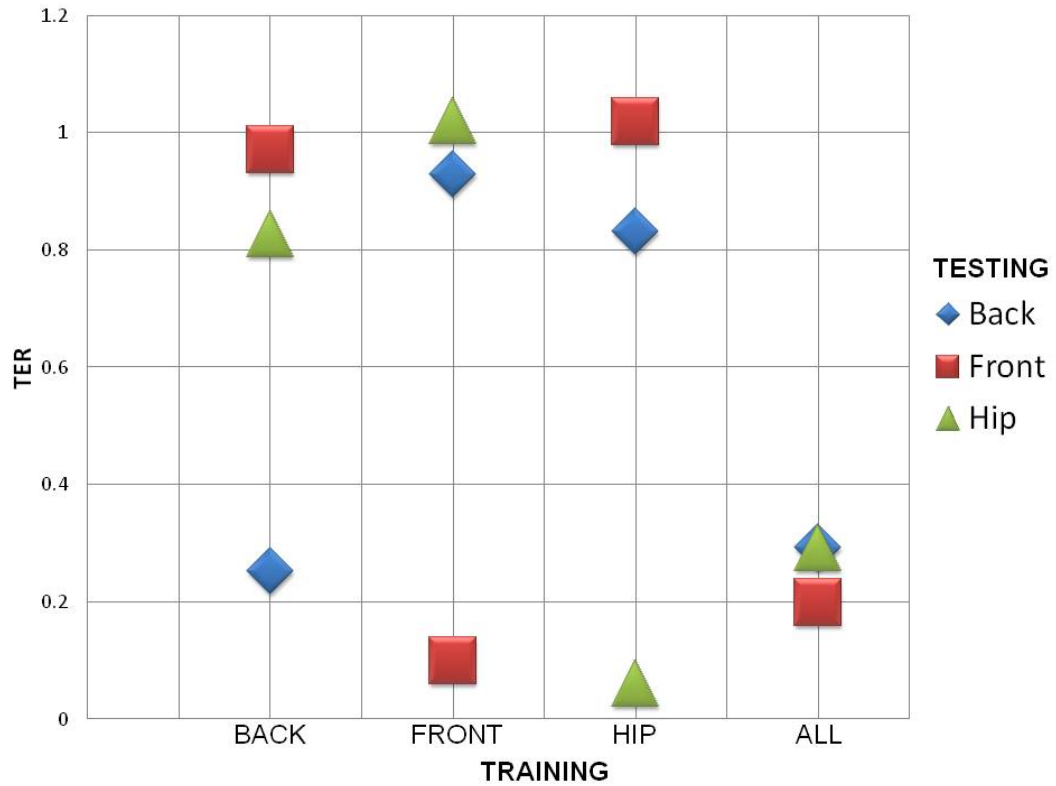
Figure 8.   Training and testing on carrying positions using zero-scaling normalization. Data from the hip position show the best performance overall.
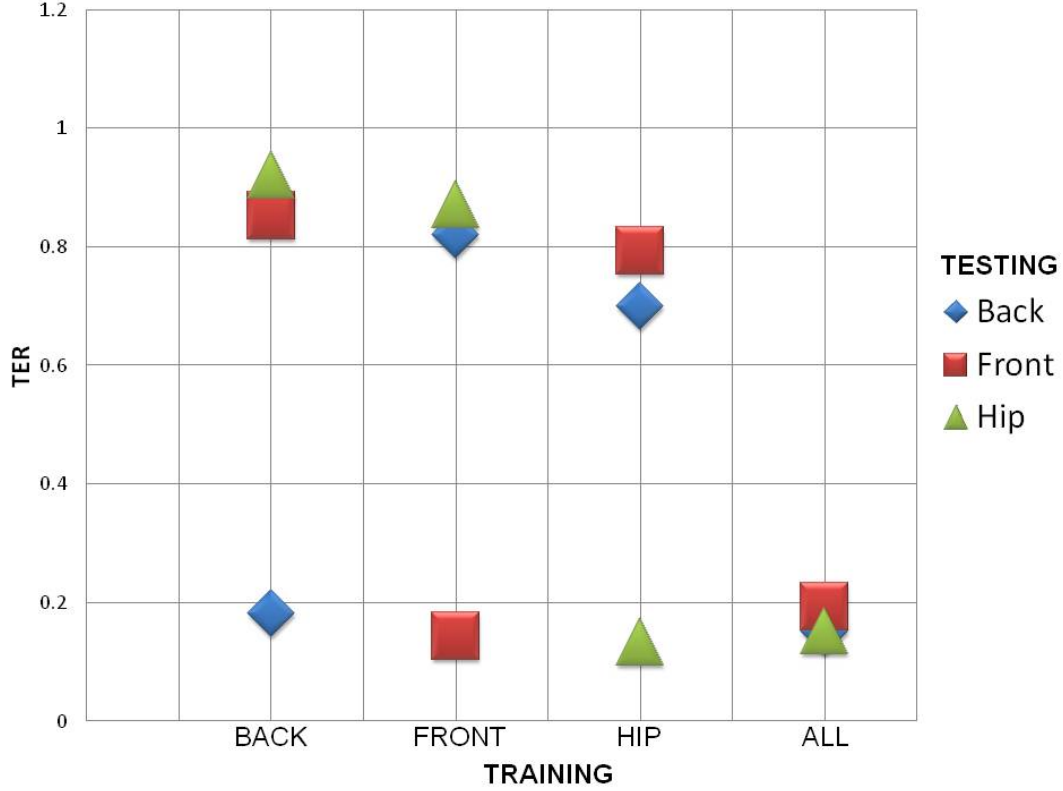
Figure 9. Training and testing on carrying positions using gravity rotation normalization. The inter-position variance decreased from the zero-scaled experiments.

C. SUMMARY OF RESULTS

After analyzing the performance of multiple combinations of processing techniques on the same data set, we found that certain features perform better. The best performing method involved first calculating the combined magnitude of the X and Y axes. This combined magnitude was then segmented and normalized around zero or normalize by a gravity rotation. A BFCC was extracted from each segment, then the segment sets were divided into two halves in order to train and test a kNN classifier, using $k = 8$.

The best result using zero-scaling normalization on a single position (hip-to-hip) was a TER of 6.13%. The best position-independent result using gravity rotation normalization was an mTER of 15.4%. The single position experiments closely resembled both Nickel [7] and Brandt [8], with TERs of 17.7% and 10.3%, respectively,

except for the data set and the selection of kNN as the classifier, thus indicating kNN as a better-performing classifier than SVM for gait authentication. In Nickel [36], kNN was also shown to be a valid classifier, though their TER of 16.48% is outperformed by our method.

THIS PAGE INTENTIONALLY LEFT BLANK

V. CONCLUSION AND RECOMMENDATIONS

This thesis presented methods of processing accelerometer-based gait signal for smartphone authentication. Our results were compared to baseline methods, as studied by others, to justify the usage of our methods in future research. Additionally, these well-performing methods were employed to evaluate the authentication performance of a system trained on data collected from multiple body positions and tested on data from unknown carrying locations.

A. OBSERVATIONS

In addition to performing a robust evaluation of the best performing classifier and axis of previous studies, we made the hypothesis that a normalization technique involving the rotation of the devices frame of reference in relation to gravity, would perform better than a zero-scaling. We also attempted to develop a gait authentication system that would not restrict a user from carrying the device in different position.

1. Rotation Performance

The effect of rotating the device's axes showed promising results. Regardless of the device's carrying position, the individual Y-axis when rotated due to gravity achieved an mTER of 12.6%. Additionally, the combined XY and XYZ axes, when rotated, achieved mTERs of 15.4% and 18.0%, respectively, indicating the dependence on the axis experiencing the strongest gravitational effect, the Y, or vertical, axis. This is better than 24.3% baseline for the Y-axis, the 19.4% for the XY-axis, and 25.7% for the XYZ-axis as well as an improvement on many previously reported results.

Data that was only zero-scaled and not rotated, as performed by other studies, showed mixed results. While the lowest position-dependent (hip-to-hip) TER of 6.13% occurred on the zero-scaled XY-axis, the data across axes had the worst mTER of all normalization techniques clustering around 33.8%. Surprisingly, data that was not normalized showed the best mTER cluster regardless of axis at 26.8%. When zero-scaling is applied after the rotation, the selection of axis is less significant as this

technique showed the lowest statistical variance across axes of all normalization techniques evaluated.

2.   Carrying Position Performance

As expected, training a classifier on a single carrying position and testing on data from another position yields poor performance. However, we were able to show that training a classifier on all carrying positions, and testing on any of the other positions, will yield consistent and satisfactory performance no matter which position the test data is captured from, particularly when the data is normalized by a gravity rotation. This important result leads us to the conclusion that any deployable smartphone-based gait authentication system should train on data captured from multiple positions, thus eliminating artificial restrictions on the placement of the device.

B.   FUTURE WORK

This thesis combined and evaluated the performance of several important settings for smartphone-based gait authentication, including axis selection, normalization techniques, and carrying position. Future work should include combining our findings with those of others, such as Nickel [7], Brandt [8], Vildjiounaite [9], Gafurov [21], and Holien [26], to determine the optimal settings when accounting for the speed of walk, types of footwear and clothing, terrain effects, and others leading to a deployable smartphone-based authentication system.

We specifically evaluated data normalization techniques of zero-scaling, gravity rotation, and gravity rotation followed by zero-scaling. The gravity rotation technique showed interesting results; however, these results may be improved through a more robust calibration technique. Our experiments collect a gravity baseline for calibration while the subject is standing still. This method seems to work well for a hip-carrying position, where the vertical axis' pitch has minimal change over the course of a walk. In the pocket positions, however, the pitch of the vertical axis is more likely to be affected by the movement of the leg. Hence, study of calibration techniques that take this leg movement into account is warranted.

In this thesis, we based our feature extraction on using the well-performing BFCC parameters described in Nickel [7] and Brandt [8]. Though these works dedicated a significant amount of study to verifying the performance of multiple sets of features, their work did not include study of devices in carrying positions aside for the hip holster. We conjecture that there may exist other signal features that may improve the classification performance for gait authentication.

We used a relatively small database of 23 subjects to evaluate the performance of our settings. Though we believe this number to be adequate to generalize our findings, research involving a larger database of subjects, collected over a period of time, could further verify our results. With a database that included other carrying positions, including in the hand and in a bag, the effect of combining multiple positions may be more fully determined. Additionally, making a larger, more robust database available publicly, would allow researchers to report comparable results for different techniques on the same data set.

Work in different areas of smartphone authentication, including Fleming [48] and Nguyen [49], should be combined with our findings in order to ensure a secure and robust system. For instance, while gait may be used to authenticate a user initially, the authentication of their typing characteristics or their wireless hotspot signature could be employed for follow-up authentication.

Once a deployable system, with multimodal authentication, is developed, evaluation of an imposter's ability to break the security of the system should be conducted.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A. VOTING PERFORMANCE

AXIS	METHOD	#V	#G	FMR	FNMR	TER
X	SVM	1	1	0.0217	0.6250	0.6467
		2	1	0.0326	0.5217	0.5543
		2	2	0.0114	0.7283	0.7397
		4	1	0.0464	0.4348	0.4812
		4	2	0.0267	0.5435	0.5702
		4	3	0.0109	0.7174	0.7283
		4	4	0.0040	0.8043	0.8083
		8	1	0.0652	0.3478	0.4130
		8	2	0.0356	0.4348	0.4704
		8	3	0.0257	0.4783	0.5040
		8	4	0.0217	0.6087	0.6304
		8	5	0.0138	0.7391	0.7529
		8	6	0.0059	0.7826	0.7885
		8	7	0.0059	0.7826	0.7885
		8	8	0.0020	0.8261	0.8281
	kNN	1	1	0.0556	0.4674	0.5230
		2	1	0.0884	0.3152	0.4036
		2	2	0.0232	0.6196	0.6428
		4	1	0.1235	0.2174	0.3409
		4	2	0.0613	0.3478	0.4091
		4	3	0.0316	0.5217	0.5533
		4	4	0.0069	0.7826	0.7895
		8	1	0.1759	0.1304	0.3063
		8	2	0.0968	0.2609	0.3577
		8	3	0.0672	0.2609	0.3281
		8	4	0.0474	0.3478	0.3952
		8	5	0.0296	0.5217	0.5513
		8	6	0.0178	0.6087	0.6265
		8	7	0.0079	0.6957	0.7036
		8	8	0.0040	0.9130	0.9170

Table 14.   Voting performance on X-axis data.

AXIS	METHOD	#V	#G	FMR	FNMR	TER
Y	SVM	1	1	0.0326	0.4239	0.4565
		2	1	0.0494	0.3043	0.3537
		2	2	0.0168	0.5435	0.5603
		4	1	0.0682	0.2391	0.3073
		4	2	0.0405	0.3043	0.3448
		4	3	0.0188	0.4348	0.4536
		4	4	0.0049	0.7174	0.7223
		8	1	0.0949	0.1739	0.2688
		8	2	0.0613	0.1739	0.2352
		8	3	0.0474	0.2174	0.2648
		8	4	0.0277	0.3478	0.3755
		8	5	0.0198	0.4348	0.4546
		8	6	0.0099	0.4783	0.4882
		8	7	0.0040	0.6087	0.6127
		8	8	0.0000	0.9565	0.9565
	kNN	1	1	0.0506	0.3967	0.4473
		2	1	0.0741	0.2717	0.3458
		2	2	0.0287	0.5217	0.5504
		4	1	0.1028	0.1957	0.2985
		4	2	0.0603	0.2826	0.3429
		4	3	0.0316	0.4130	0.4446
		4	4	0.0109	0.6957	0.7066
		8	1	0.1383	0.1739	0.3122
		8	2	0.0810	0.1739	0.2549
		8	3	0.0692	0.1739	0.2431
		8	4	0.0534	0.2174	0.2708
		8	5	0.0316	0.4783	0.5099
		8	6	0.0198	0.5217	0.5415
		8	7	0.0138	0.6087	0.6225
		8	8	0.0040	0.8261	0.8301

Table 15.   Voting performance on Y-axis data.

AXIS	METHOD	#V	#G	FMR	FNMR	TER
Z	SVM	1	1	0.0440	0.4837	0.5277
		2	1	0.0623	0.3913	0.4536
		2	2	0.0272	0.5761	0.6033
		4	1	0.0800	0.3478	0.4278
		4	2	0.0514	0.3696	0.4210
		4	3	0.0326	0.5000	0.5326
		4	4	0.0148	0.7174	0.7322
		8	1	0.0968	0.3043	0.4011
		8	2	0.0711	0.3043	0.3754
		8	3	0.0593	0.3478	0.4071
		8	4	0.0455	0.3913	0.4368
		8	5	0.0375	0.4348	0.4723
		8	6	0.0277	0.5217	0.5494
		8	7	0.0138	0.6957	0.7095
		8	8	0.0059	0.8696	0.8755
	kNN	1	1	0.0679	0.4402	0.5081
		2	1	0.0963	0.3478	0.4441
		2	2	0.0420	0.5326	0.5746
		4	1	0.1314	0.2826	0.4140
		4	2	0.0771	0.3913	0.4684
		4	3	0.0435	0.4348	0.4783
		4	4	0.0247	0.6522	0.6769
		8	1	0.1739	0.2609	0.4348
		8	2	0.1146	0.3043	0.4189
		8	3	0.0791	0.3478	0.4269
		8	4	0.0573	0.3913	0.4486
		8	5	0.0474	0.3913	0.4387
		8	6	0.0415	0.4348	0.4763
		8	7	0.0237	0.6087	0.6324
		8	8	0.0158	0.7826	0.7984

Table 16.   Voting performance on Z-axis data.

AXIS	METHOD	#V	#G	FMR	FNMR	TER
XY	SVM	1	1	0.0292	0.4674	0.4966
		2	1	0.0459	0.3098	0.3557
		2	2	0.0126	0.6250	0.6376
		4	1	0.0692	0.1957	0.2649
		4	2	0.0331	0.3261	0.3592
		4	3	0.0104	0.5543	0.5647
		4	4	0.0044	0.7935	0.7979
		8	1	0.0988	0.1087	0.2075
		8	2	0.0593	0.1739	0.2332
		8	3	0.0356	0.2391	0.2747
		8	4	0.0168	0.3478	0.3646
		8	5	0.0109	0.5000	0.5109
		8	6	0.0079	0.6739	0.6818
		8	7	0.0040	0.8043	0.8083
		8	8	0.0009	0.8913	0.8923
	kNN	1	1	0.0463	0.3804	0.4267
		2	1	0.0721	0.2228	0.2949
		2	2	0.0208	0.5380	0.5588
		4	1	0.1042	0.1196	0.2238
		4	2	0.0509	0.2283	0.2792
		4	3	0.0227	0.4348	0.4575
		4	4	0.0079	0.7391	0.7470
		8	1	0.1462	0.0870	0.2332
		8	2	0.0850	0.1087	0.1937
		8	3	0.0573	0.1522	0.2095
		8	4	0.0346	0.1957	0.2303
		8	5	0.0198	0.3696	0.3894
		8	6	0.0158	0.5435	0.5593
		8	7	0.0079	0.6957	0.7036
		8	8	0.0049	0.8913	0.8962

Table 17.   Voting performance on XY-axis data.

AXIS	METHOD	#V	#G	FMR	FNMR	TER
XYZ	SVM	1	1	0.0341	0.4837	0.5178
		2	1	0.0504	0.3587	0.4091
		2	2	0.0183	0.6087	0.6270
		4	1	0.0702	0.2391	0.3093
		4	2	0.0415	0.4130	0.4545
		4	3	0.0178	0.5652	0.5830
		4	4	0.0079	0.7174	0.7253
		8	1	0.0929	0.1739	0.2668
		8	2	0.0613	0.2609	0.3222
		8	3	0.0455	0.3478	0.3933
		8	4	0.0277	0.3478	0.3755
		8	5	0.0198	0.5217	0.5415
		8	6	0.0138	0.6957	0.7095
		8	7	0.0099	0.7391	0.7490
		8	8	0.0040	0.7826	0.7866
	kNN	1	1	0.0553	0.3750	0.4303
		2	1	0.0805	0.2500	0.3305
		2	2	0.0306	0.5000	0.5306
		4	1	0.1107	0.1739	0.2846
		4	2	0.0573	0.2826	0.3399
		4	3	0.0405	0.3696	0.4101
		4	4	0.0138	0.6739	0.6877
		8	1	0.1581	0.1304	0.2885
		8	2	0.0830	0.1739	0.2569
		8	3	0.0672	0.2174	0.2846
		8	4	0.0494	0.2609	0.3103
		8	5	0.0415	0.3478	0.3893
		8	6	0.0257	0.4783	0.5040
		8	7	0.0119	0.6087	0.6206
		8	8	0.0079	0.7826	0.7905

Table 18. Voting performance on XYZ-axis data.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B. CLASSIFIER PERFORMANCE

AXIS	NORMALIZATION	kNN			SVM		
		TER	FMR	FNMR	TER	FMR	FNMR
X	None	0.3280	0.0237	0.3043	0.5435	0.0217	0.5217
	Zerod	0.3774	0.0296	0.3478	0.4150	0.0237	0.3913
	Rotated	0.3814	0.0336	0.3478	0.4111	0.0198	0.3913
	Rotated Zerod	0.3814	0.0336	0.3478	0.3260	0.0217	0.3043
Y	None	0.2411	0.0237	0.2174	0.2273	0.0099	0.2174
	Zerod	0.6600	0.3557	0.3043	0.5729	0.0079	0.5650
	Rotated	0.1561	0.0257	0.1304	0.2371	0.0198	0.2174
	Rotated Zerod	0.3794	0.0316	0.3478	0.4941	0.0158	0.4783
Z	None	0.3320	0.0277	0.3043	0.3952	0.0039	0.3913
	Zerod	0.3320	0.0277	0.3043	0.5415	0.0198	0.5217
	Rotated	0.4249	0.0336	0.3913	0.4130	0.0217	0.3913
	Rotated Zerod	0.3814	0.0336	0.3478	0.5830	0.0178	0.5652
XY	None	0.1048	0.0178	0.0870	0.1838	0.0099	0.1739
	Zerod	0.2530	0.0356	0.2174	0.3122	0.0079	0.3043
	Rotated	0.1521	0.0217	0.1304	0.2352	0.0178	0.2174
	Rotated Zerod	0.1877	0.0138	0.1739	0.4426	0.0079	0.4347
XYZ	None	0.1047	0.0178	0.0869	0.2371	0.0198	0.2174
	Zerod	0.3339	0.0296	0.3043	0.4506	0.0158	0.4348
	Rotated	0.1047	0.0178	0.0869	0.2371	0.0198	0.2174
	Rotated Zerod	0.1205	0.0336	0.0869	0.4091	0.0178	0.3913

Table 19.   kNN and SVM results in back pocket carrying position.

AXIS	NORMALIZATION	kNN			SVM		
		TER	FMR	FNMR	TER	FMR	FNMR
X	None	0.4684	0.0336	0.4348	0.5000	0.0217	0.4782
	Zerod	0.4288	0.0375	0.3913	0.5395	0.0178	0.5217
	Rotated	0.4249	0.0336	0.3913	0.4938	0.0158	0.4780
	Rotated Zerod	0.2075	0.0336	0.1739	0.3241	0.0198	0.3043
Y	None	0.2035	0.0296	0.1739	0.4051	0.0138	0.3913
	Zerod	0.3750	0.0271	0.3478	0.4505	0.0158	0.4347
	Rotated	0.1106	0.0237	0.0869	0.3597	0.0119	0.3478
	Rotated Zerod	0.2450	0.0277	0.2174	0.4071	0.0158	0.3913
Z	None	0.2411	0.0237	0.2174	0.2806	0.0198	0.2609
	Zerod	0.3854	0.0376	0.3478	0.3181	0.0138	0.3043
	Rotated	0.3419	0.0376	0.3043	0.4190	0.0277	0.3913
	Rotated Zerod	0.2945	0.0336	0.2609	0.3695	0.0217	0.3478
XY	None	0.1976	0.0237	0.1739	0.2312	0.0138	0.2174
	Zerod	0.0988	0.0119	0.0869	0.1363	0.0059	0.1304
	Rotated	0.1996	0.0257	0.1739	0.3162	0.0119	0.3043
	Rotated Zerod	0.3320	0.0277	0.3043	0.4467	0.0119	0.4348
XYZ	None	0.2371	0.0198	0.2174	0.3636	0.0158	0.3478
	Zerod	0.2075	0.0336	0.1739	0.3597	0.0119	0.3478
	Rotated	0.2371	0.0198	0.2174	0.3636	0.0158	0.3478
	Rotated Zerod	0.4209	0.0296	0.3913	0.6304	0.0217	0.6087

Table 20.   kNN and SVM results in front pocket carrying position.

AXIS	NORMALIZATION	kNN			SVM		
		TER	FMR	FNMR	TER	FMR	FNMR
X	None	0.5119	0.0336	0.4782	0.5020	0.0237	0.4783
	Zerod	0.4269	0.0355	0.3913	0.4051	0.0138	0.3913
	Rotated	0.6561	0.0474	0.6087	0.7233	0.0277	0.6957
	Rotated Zerod	0.3854	0.0376	0.3478	0.5830	0.0178	0.5652
Y	None	0.2016	0.0277	0.1739	0.3162	0.0119	0.3043
	Zerod	0.1620	0.0316	0.1304	0.2747	0.0138	0.2609
	Rotated	0.1107	0.0237	0.0870	0.2786	0.0178	0.2609
	Rotated Zerod	0.1542	0.0237	0.1304	0.2312	0.0138	0.2174
Z	None	0.3735	0.0257	0.3478	0.4525	0.0178	0.4347
	Zerod	0.5454	0.0237	0.5217	0.4506	0.0158	0.4348
	Rotated	0.4289	0.0376	0.3913	0.5415	0.0198	0.5217
	Rotated Zerod	0.2945	0.0336	0.2609	0.4426	0.0079	0.4347
XY	None	0.2016	0.0277	0.1739	0.3616	0.0138	0.3478
	Zerod	0.0613	0.0178	0.0435	0.3201	0.0158	0.3043
	Rotated	0.1106	0.0237	0.0869	0.1423	0.0119	0.1304
	Rotated Zerod	0.2510	0.0336	0.2174	0.3241	0.0198	0.3043
XYZ	None	0.1996	0.0257	0.1739	0.2747	0.0138	0.2609
	Zerod	0.4170	0.0267	0.3913	0.4980	0.0198	0.4783
	Rotated	0.1996	0.0257	0.1739	0.2746	0.0138	0.2608
	Rotated Zerod	0.2925	0.0316	0.2609	0.3992	0.0079	0.3913

Table 21.   kNN and SVM results in hip holster carrying position.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX C. POSITION-INDEPENDENT PERFORMANCE

AXIS	NORMALIZATION	mTER	VARIANCE
X	None	0.4361	0.0062
	Zero-Scale	0.4110	0.0006
	Rotated	0.4874	0.0145
	Rotated and Zero	0.3247	0.0069
Y	None	0.2154	0.0003
	Zero-Scale	0.3990	0.0416
	Rotated	0.1258	0.0005
	Rotated and Zero	0.2595	0.0086
Z	None	0.3155	0.0031
	Zero-Scale	0.4209	0.0082
	Rotated	0.3985	0.0016
	Rotated and Zero	0.3234	0.0017
XY	None	0.1680	0.0031
	Zero-Scale	0.1377	0.0074
	Rotated	0.1541	0.0031
	Rotated and Zero	0.2569	0.0151
XYZ	None	0.1805	0.0020
	Zero-Scale	0.3195	0.0069
	Rotated	0.1805	0.0013
	Rotated and Zero	0.2780	0.0035

Table 22.   mTER and inter-position variance of axis-normalization mixtures across positions.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX D. NORMALIZATION PERFORMANCE PER POSITION

A.   NORMALIZATION IN BACK POCKET

	NONE	ZERO-SCALED	ROTATED	ROTATED AND ZERO-SCALED
X	0.3280	0.3774	0.3814	0.3814
Y	0.2411	0.6600	0.1561	0.3794
Z	0.3320	0.3320	0.4249	0.3814
XY	0.1048	0.2530	0.1521	0.1877
XYZ	0.1048	0.3339	0.1048	0.1205
Mean	0.2221	0.3913	0.2438	0.2901
Variance	0.0102	0.0197	0.0174	0.0128

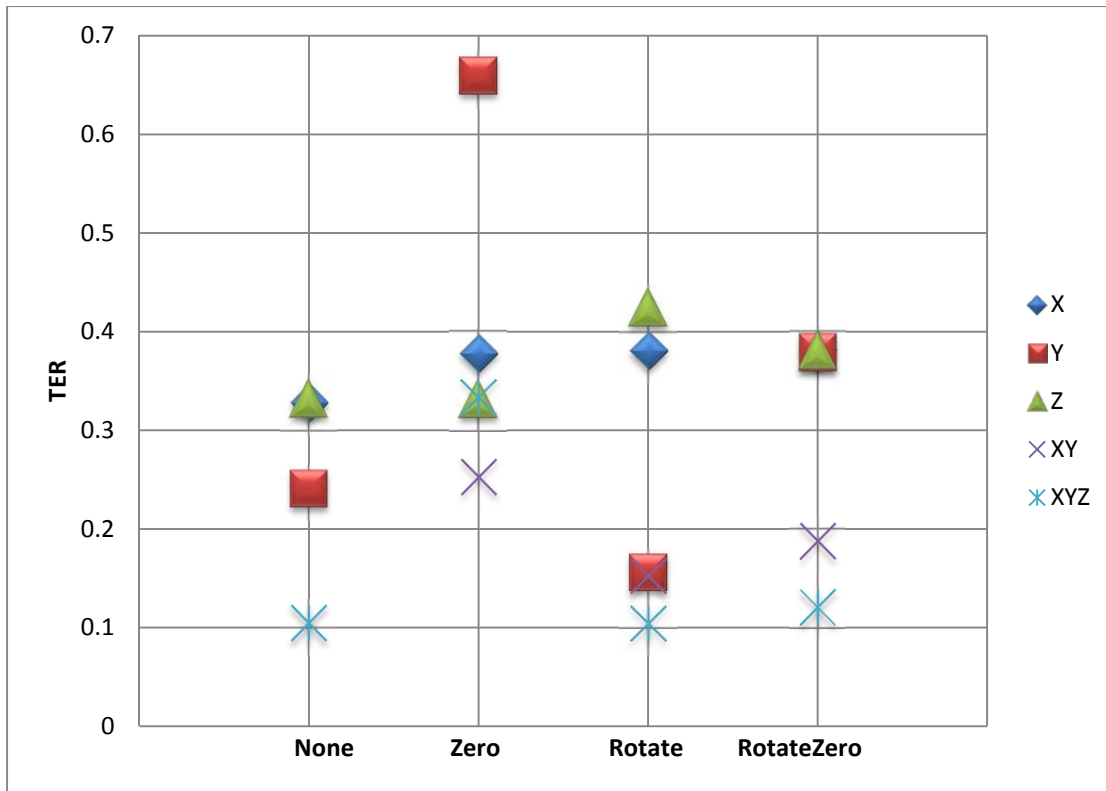Table 23.   Performance of normalization techniques in back pocket carrying position.

Figure 10.   Performance of normalization techniques in back pocket carrying position.

B.   NORMALIZATION IN FRONT POCKET

	NONE	ZERO-SCALED	ROTATED	ROTATED AND ZERO-SCALED
X	0.4684	0.4288	0.4249	0.2075
Y	0.2035	0.3749	0.1106	0.2451
Z	0.2411	0.3853	0.3418	0.2945
XY	0.1976	0.0988	0.1996	0.3320
XYZ	0.2372	0.2075	0.2371	0.4209
Mean	0.2696	0.2991	0.2628	0.3000
Variance	0.0102	0.0157	0.0121	0.0055

Table 24.   Performance of normalization techniques in front pocket carrying position.
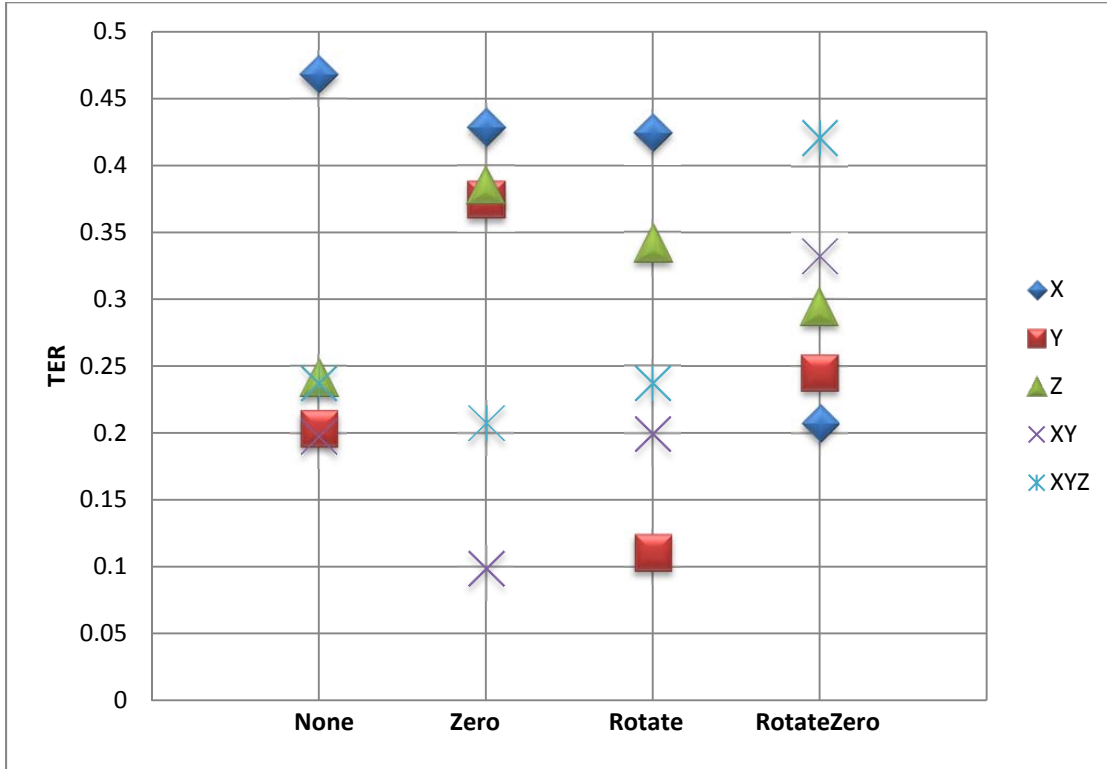

Figure 11.   Performance of normalization techniques in front pocket carrying position.

C. NORMALIZATION IN HIP HOLSTER

	NONE	ZERO-SCALED	ROTATED	ROTATED AND ZERO-SCALED
X	0.5119	0.4269	0.6561	0.3854
Y	0.2016	0.1620	0.1106	0.1541
Z	0.3735	0.5454	0.4288	0.2945
XY	0.2016	0.0613	0.1106	0.2510
XYZ	0.1996	0.4170	0.1996	0.2925
Mean	0.2976	0.3225	0.3011	0.2755
Variance	0.0159	0.0327	0.0450	0.0056

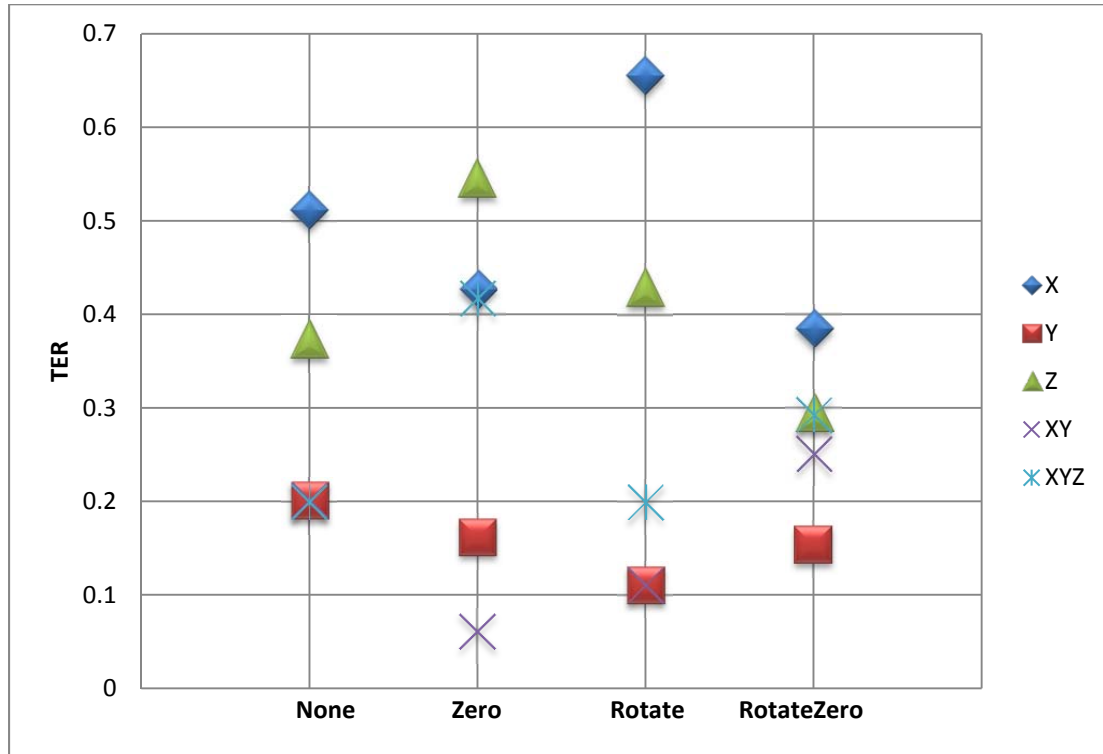Table 25.   Performance of normalization techniques in hip carrying position.

Figure 12.   Performance of normalization techniques in hip carrying position.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

[1]   Ponemon Research Institute LLC., "The lost smartphone problem: Benchmark study of U.S. organizations," *Ponemon Institute Research Report* sponsored by McAfee, Traverse City, MI, Oct 2011.

[2]   Federal Communications Commission (2012, April 10). Announcement of new initiatives to combat smartphone and data theft. [Online]. Available: http://www.fcc.gov/document/announcement-new-initiatives-combat-smartphone-and-data-theft

[3]   Ponemon Research Institute LLC., "Global study on mobility risks," *Ponemon Institute Research Report* sponsored by Websense Inc, Traverse City, MI, Feb 2012.

[4]   I. Arghire. (2009, July 10). Accelerometers in one of three phones by 2010. Softpedia [Online]. Available: http://news.softpedia.com/news/Accelerometers-in-One-of-Three-Phones-by-2010–116360.shtml

[5]   K. Bostic. (2013, May 13). New York authorities ask Apple, Google to help stop smartphone thefts. Apple Insider [Online]. Available: http://appleinsider.com/articles/13/05/13/new-york-authorities-ask-apple-google-to-help-stop-smartphone-thefts

[6]   D. Gafurov *et al.,* "Biometric gait authentication using accelerometer sensor," *Journal of Computers*, vol. 1, no.7, pp. 51–59, Nov 2006.

[7]   C. Nickel, "Accelerometer-based biometric gait recognition for authentication on smartphones," Ph. D. dissertation, Dept. of Comp. Sci, Darmstadt Univ of Technology, Darmstadt, DE, 2012.

[8]   H. Brandt, "Classification of acceleration data for biometric gait recognition on mobile devices," M.S. thesis, Dept. of Comp. Sci., Hochschule Darmstadt Univ. of App. Sci, Darmstadt, DE, 2010.

[9]   E. Vildjiounaite *et al.*, "Unobtrusive multimodal biometrics for ensuring privacy and information security with personal devices," *Pervasive Computing*, K. Fishkin, B. Schiele, P. Nixon, and A. Quigley, Eds, Berlin, DE: Springer Berlin Heidelberg, *Lecture Notes in Computer Science*, vol. 3968, pp. 187–201. 2006

[10]   L. Bianchi *et al.*, "Individual characteristics of human walking mechanics," *Pflugers Archiv: European Journal of Physiology*, vol. 436, no. 3, pp. 343–356, June 1998.

[11]   *Biometrics Metrics Report*, ver 3, U.S. Military Academy, West Point, NY, 2012.

[12] M. Bromba. (2012, Nov 2). BioIdentification [Online]. Available: http://www.bromba.com/faq/biofaqe.htm.

[13] M. S. Nixon *et al.*, "Automatic gait recognition," *Biometrics*, A.K. Jain, R. Bolle, and S. Pankanti, Eds. New York, NY: Springer USA, ch. 11, pp. 231–249, 1996.

[14] J. Han and B. Bhanu, "Individual recognition using gait energy image," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 28, no. 2, pp. 316–322, Feb. 2006.

[15] Z. Liu and S. Sarkar, "Improved gait recognition by gait dynamics normalization," *IEEE Trans. on Pattern Analysis and Machine Intelligence*," vol. 28, no. 6, pp. 863–876, June 2006.

[16] K. Nakajima *et al.*, "Footprint-based personal recognition," *IEEE Trans. on Biomedical Engineering*, vol. 47, no. 11, pp. 1534–1537, Nov. 2000.

[17] J. Jenkins and C. Ellis, "Using ground reaction forces from gait analysis: body mass as a weak biometric," *Pervasive Computing*, A. LaMarca, M. Langheinrich, and K. Truong, Eds, Berline, DE: Springer Berline Heidelberg, *Lecture Notes in Computer Science*, vol. 4480, pp. 251–267, 2007.

[18] H. A. Ailisto *et al.*, "Identifying people from gait pattern with accelerometers," *Proc Biometric Technology for Human Identification II*, Orlando, FL, 2005

[19] J. Mäntyjärvi *et al.*, "Identifying users of portable devices from gait pattern with accelerometers," *IEEE Int Conf on Acoustics, Speech, and Signal Processing, Proc (ICASSP '05)*, vol. 2, pp.ii/973-ii/976, March 2005.

[20] D. Gafurov *et al.*, "Robustness of biometric gait authentication against impersonation attack," *On the Move to Meaningful Internet Systems*, R. Meersman, Z. Tari, and P. Herrero, Eds. Berlin, DE: Springer Berlin Heidelberg, *Lecture Notes in Computer Science*, vol. 4277, pp. 479–488, 2006.

[21] D. Gafurov *et al.*, "Gait authentication and identification using wearable accelerometer sensors," *2007 IEEE Workshop on Automatic Identification Advanced Technology*, pp. 220–225, June 2007.

[22] S. Sprager and D. Zazula, "A cumulant-based method for gait identification using accelerometer data with principal component analysis and support vector machine," *WSEAS Trans on Signal Processing*, vol. 5, no.11, pp. 369–378, 2009.

[23] L. Rong *et al.*, "Identification of individual walking patterns using gait acceleration," *The 1st Int Conf on Bioinformatics and Biomedical Engineering*, pp. 543–546, July 2007.

[24]   M. O. Derawi *et al.*, "Unobtrusive user-authentication on mobile phones using biometric gait recognition," *2010 Sixth Int Conf on Intelligent Information hiding and Multimedia Signal Processing (IIH-MSP)*, pp.306–311. Oct 2010.

[25]   D. Gafurov *et al.*, "Improved gait recognition performance using cycle matching," *2010 IEEE 24th Int Conf on Advanced Information Networking and Applications Workshops (WAINA)*, pp. 836–841, April 2010.

[26]   K. Holien, "Gait recognition under non-standard circumstances," M.S. thesis, Dept. of Comp. Sci. and Media Tech., Gjøvik University College, Gjøvik, NO, 2008

[27]   B. P Bogert *et al.*, "The quefrency analysis of time series for echoes: cepstrum, pseudoautocovariance, cross-cepstrum and saphe cracking," *Proc of Symposium Time Series Analysis*, 1963

[28]   A. V. Oppenheim and R. W. Schafer, "From frequency to quefrency: a history of the cepstrum," *IEEE Signal Processing Magazine*, vol. 21, no. 5, pp. 95–106, Sept 2004.

[29]   P. Mermelstein, "Distance measures for speech recognition, psychological and instrumental," *Pattern Recognition and Artificial Intelligence*, vol. 116, pp.374–388, 1976.

[30]   S. Davis and P. Mermelstein, "Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences," *IEEE Trans on Acoustic, Speech and Signal Processing*, vol. 28, no.4, pp.357–366, Aug 1980.

[31]   J. R. Kwapisz *et al.*, "Activity recognition using cell phone accelerometers," *ACM SIGKDD Explorations Newsletter*, vol. 12, no. 2, pp. 74–82, Dec 2010.

[32]   T. Zhang *et al.*, "Fall detection by embedding an accelerometer in cellphone and using KFD algorithm." *Int Journal of Computer Science and Network Security,* vol. 6, no. 10, pp. 277–284, 2006.

[33]   J. Dai *et al.*, "PerFallD: a pervasive fall detection system using mobile phones," *2010 8th IEEE Int Conf on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, pp. 292–297, April 2010.

[34]   M. D. Tundo *et al.*, "Correcting smartphone orientation for accelerometer-based analysis," *2013 IEEE Int Symp on Medical Measurements and Applications Proc (MeMeA)*, pp. 58–62, May 2013.

[35]   J. P. Lewis, "A short SVM (support vector machine) tutorial," University of Southern California, CGIT Lab/IMSC, ver. 0.zz, Dec 2004.

[36]   C. Nickel *et al.*, "Authentication of smartphone users based on the way they walk using k-NN algorithm," *2012 Eighth Int Conf on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP)*, pp. 16–20, July 2012.

[37]   J. Demšar *et al.*, "Orange: data mining toolbox in Python," *Journal of Machine Learning Research*, vol. 17, pp. 2349–2353, April 2013.

[38]   Mind Treat Studios. (2012, June 26). "How to determine which hand holds the smartphone using its accelerometer." [Online]. Available: http://www.mindtreatstudios.com/how-its-made/kinect-determine-hand-holds-smartphone-accelerometer/

[39]   T. E. Oliphant, "Python for scientific computing," *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90, 2007.

[40]   J. Cooke *et al.*, "Flight simulation dynamic modeling using quaternions," *Presence*, vol. 1, no. 4, pp. 404–420, 1992.

[41]   E. B. Dam *et al.*, "Quaternions, interpolation, and animation," Dept. of Comp. Sci., Univ. of Copenhagen, Denmark, Tech. Rep. 98/5, July 1998.

[42]   E. Zwicker, "Subdivision of the audible frequency range into critical bands (frequenzgruppen)," *The Journal of the Acoustical Society of America*, vol. 33, no. 2, pp. 248, 1961.

[43]   L. Rabiner and B. H. Juang, *Fundamentals of Speech Recognition*, Upper Saddle River, NJ: Prentice-Hall, 1993.

[44]   H. Traunmüller, "Analytical expressions for the tonotopic sensory scale," *The Journal of the Acoustical Society of America*, vol. 88, no. 1, pp. 97–100, 1990.

[45]   B. Logan, "Mel frequency cepstral coefficients for music modeling," *Int Symp on Music Information Retreival*, 2000.

[46]   D. Ellis. (2005). PLP and RASTA (and MFCC, and inversion) in MATLAB. [Online]. Available: http://www.ee.columbia.edu/~dpwe/resources/matlab/rastamat/

[47]   C. Hsu *et al.* (2003). A practical guide to support vector classification. [Online] Available: http://www.cs.sfu.ca/people/Faculty/teaching/726/spring11/svmguide.pdf

[48]   S. Fleming, "Identification of a smartphone user via keystroke analysis," M.S. thesis, Dept. of Comp. Sci., Naval Postgraduate School, Monterey, CA, 2014.

[49]   V. Nguyen, "Continuous authentication of a smartphone user using RSSI-based geolocation," M.S. thesis, Dept. of Comp. Sci., Naval Postgraduate School, Monterey, CA, 2014.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
    Ft. Belvoir, Virginia

2. Dudley Knox Library
    Naval Postgraduate School
    Monterey, California